

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: January 17, 2004, 01:15:10 ; Search time 10931 Seconds

(without alignments)  
5344.337 Million cell updates/sec

Title: US-10-020-095-4

Perfect score: 7348

Sequence: 1 MGGPPLTFAHLHLCVCTAAL.....HSSVIFIFCFKLYFMEIWL 1428

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DBV=rlp  
-Q=/cgn2.1/USPTO.spool\_p/US10020095/runat\_15012004\_153822\_18707/app\_query.fasta\_1.1607  
-DB=GenEmbl -QFMT=fastap -SUFFIX=ige -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10020095.@CCN\_1\_1\_7794@runat\_15012004\_153822\_18707 -NCPU=6 -ICPU=3  
-NO\_MMAPP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
1: GenEmbl:\*  
2: gb\_da:\*  
3: gb\_htg:\*  
4: gb\_in:\*  
5: gb\_om:\*  
6: gb\_ov:\*  
7: gb\_pat:\*  
8: gb\_ph:\*  
9: gb\_pl:\*  
10: gb\_pr:\*  
11: gb\_ro:\*  
12: gb\_sets:\*  
13: gb\_sy:\*  
14: gb\_un:\*  
15: gb\_vl:\*  
16: em\_ba:\*  
17: em\_fun:\*  
18: em\_hum:\*  
19: em\_in:\*  
20: em\_mu:\*  
21: em\_om:\*  
22: em\_or:\*  
23: em\_ov:\*  
24: em\_pac:\*  
25: em\_ph:\*  
26: em\_pl:\*  
27: em\_ro:\*  
28: em\_stg:\*  
29: em\_un:\*

## SUMMARY

SEQ ID NO: 3  
SEQ ID NO: 4  
REV. TRANS. SEQ ID NO: 4

29: em\_vl:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pln:\*  
35: em\_htg\_rtd:\*  
36: em\_htg\_mam:\*  
37: em\_htg\_vrt:\*  
38: em\_htg:\*  
39: em\_htgo\_hum:\*  
40: em\_htgo\_mus:\*  
41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7326.5	99.7	4335	6	AX537610 Sequence
2	7326.5	99.7	4761	6	AX534934 Sequence
3	7326.5	99.7	5883	9	AF410459 Homo sapi
4	7326.5	99.7	5895	6	AX534938 Sequence
5	7317.5	99.6	4335	6	AX537612 Sequence
6	7317.5	99.6	4761	6	AX534936 Sequence
7	7317.5	99.6	5895	6	AX534940 Sequence
8	5470	74.4	3535	6	AX534942 Sequence
9	5461	74.3	3535	6	AX534944 Sequence
10	5360	72.9	4735	10	AY083458 Mus muscu
11	5360	72.9	5644	10	BC052443 Mus muscu
12	4305.5	56.6	2938	9	HSN80581
13	3360	45.7	2273	9	AK095888 Homo sapi
14	1854.5	25.2	4527	6	AX054953 Sequence
15	1850	25.2	4560	6	AX054954 Sequence
16	1779.5	24.2	5603	3	CJN431688 Ciona int
17	1692	22.0	4578	3	AY122084 Drosophi
18	1692	22.0	4579	3	DME269539 Drosophi
19	1638	22.3	4678	3	DB3196
20	1542.5	21.0	5817	9	AB033109 Homo sapi
21	1542.5	21.0	5852	9	AY101765 Homo sapi
22	1491	20.3	4545	10	GPIAMA
23	1481.5	20.2	5210	9	HSN803446
24	1463.5	19.9	4598	3	DME269541
25	1460	19.9	4220	3	DME269538
26	1459	19.9	4944	3	AF538967
27	1455	19.8	4636	10	RAT113A
28	1441	19.6	4577	6	AR241630
29	1440	19.6	4425	6	AX054943 Sequence
30	1440	19.6	4577	6	AX146876 Sequence
31	1440	19.6	4577	6	AX409589 Sequence
32	1440	19.6	4577	9	HMA2M
33	1440	19.6	4617	9	BC040071 Homo sapi
34	1440	19.6	4809	6	AX281769 Sequence
35	1437	19.6	4801	5	HSN804792
36	1435.5	19.5	4841	5	LARA2M
37	1434	19.5	4569	6	A21185
38	1432.5	19.5	4622	10	AY185125 Mus muscu
39	1428.5	19.4	4637	3	AB050668 Branchios
40	1425.5	19.4	4650	3	DME269540 Drosophi
41	1425.5	19.4	4913	3	AY069050 Drosophi
42	1422	19.4	4488	6	AX497234 Sequence
43	1422	19.4	4488	6	AX574598 Sequence
44	1421	19.3	4599	6	A21187
45	1420	19.3	4672	10	BC051037 Mus muscu

RESULT 1

## ALIGNMENTS

AX537610  
 LOCUS AX537610 4335 bp DNA linear PAT 23-NOV-2002  
 DEFINITION Sequence 1 from Patent WO02070738.  
 ACCESSION AX537610  
 VERSION AX537610.1 GI:25269531  
 KEYWORDS  
 SOURCE  
 ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1  
 AUTHORS Schuh, A. and Ouehand, W.  
 TITLE Diagnosis and treatment of blood disorders.  
 JOURNAL Patent: WO 02070738-A 1 12-SEP-2002;  
 Schuh, Andre (CA)  
 FEATURES  
 SOURCE Location/Qualifiers  
 1. 4335  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 1. .>4335  
 /note="unamed protein product"  
 /codon\_start=1  
 /protein\_id="CAD57496.1"  
 /db\_xref="GI:25269532"  
 /translation="MQGPPLTLTAHLCTCTALAVAGPRFLVTADGIIIRPGNVTI  
 GVLELHCPQVTVKALLKTSANLTVSLAEVGFPGSKFTLLPSLPLNSADEY  
 ELRYGRDDELFSNSTRLSFETKRISVFIPTDKALYKPOEYKFRITVLSPFKRY  
 KTSILNLIKDKSKMLIQOMLSQSDLGIVISKTFLPSHPIIDGMSIOVQNDQYYS  
 FQSEVYLPKREVTLPVLYCSMNSKHLNGITAKYVYKVKADVTLLTVESVTGI  
 KKNITTKFKINGSANFSNDEEMKNVDSNGLSEYLDSPGVEILITVTVESVTGI  
 SRNVSTVFQKHDIYIEFDYTVLKLPSLFTAVKTVTRADQNLLEBRNNVIT  
 VTORNYEYMGSGNOKMAVOKINVTYPOSGTKIPEPILEDSELDLQKAYFGS  
 KSSMAVSLFKPSKTYIOLKTRDENTKVGSPRELVYSGNKLKEISYMTVSGQIVAV  
 VGKQNTMFSITPENSTPRACVYVYIEDGEIISVLYKIPVQIVFRNKIKLYMSKY  
 KAPSEKVSLEKISVQPDSTIVGIVAVKSVNMAASNDITMENVAHELELYNTGYLYG  
 MFNMSFAVFOECGLMVLIDANLTDYIDGVYDNAEYARFMEENBEGHIVIHPSLGS  
 SPYRKHPEPTMILWDTNMGYRIYQEEFVYVDSITSMVATGFIISDLGIKTTTVY  
 ELQAPQPEPFIPLNLYSVIRGESEFALETIIPNYLKADTEVYVIEKSDKDIIMTSB  
 INATGHOOTLILVSEDAATVLPRIKPHLITPTVVALSPETASADTOMIYKABEI  
 EKYSQSLIDLTNNRLOSTIKTISFSPRPPTVTGSEVQITAGDVLGFSINGLASL  
 IRMPYGGCEOMINIPAPVYIYLDYTKKQITDNLKKAALSFMKQYQRELLVQREGL  
 SFSAFQNDPESGTSWLSAFVLCFLPADPYIDQNVLHRTYMLKGHQSGNFMDFP  
 GSVHSELOGKNSPVTLAYIVTSLGKYKQPNIDVQESIHLESEFSGISDNYT  
 LALITVALSSVSGKAKALNMLTWRBOEGQMOWFWSSESLSDSQPSLDEYVA  
 YALISHLOPOTSGIPIKRWLSORNSIGSPASTODTVALKALSFEALMTERPN  
 IQVTVTGSPSPVYKFLIDTNRLDLQTAELAVQMAVINSANGEPATICQLVNVTN  
 VKAGSSRRRRRSIQNOEAFDLDAVKENKDLNHDNLNCTSFSGPERSGALMEVNL  
 LSGPMVSEALISSETVYKVEDHKLNLVNEOTFCVNIIPAVENFVSNTOQDS  
 VSIYDYEPREAOAVRSYNSVKLSSCDLVQCRPCEDDASGSHHSVIFIFCPK  
 LLYFMELML"  
 BASE COUNT 1296 a 837 c 922 g 1280 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 0 Length: 4335  
 Score: 7326.50 Matches: 1427  
 Percent Similarity: 98.82% Conservative: 1  
 Best Local Similarity: 98.75% Mismatches: 0  
 Query Match: 99.71% Indels: 17  
 Gaps: 1  
 US-10-020-095-4 (1-1428) x AX537610 (1-4335)  
 QY 1 MetGlnGlyProProLeuLeuThrAlaAlaHisLeuLeuCyValCyThrAlaAlaLeu 20  
 Db 1 ATGAGAGGAGCCACGCTCTGACCGCGCCACCTCTCTCGTGTGACGCGCGCGCTG 60  
 QY 21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGlyGly 40  
 Db 61 GCCCTGGCTCCCGGCGCTTCTGTGTGACAGCCCGAGAACTATCAGGCGCGGAGGA 120  
 QY 41 AsnValThrIleGlyValGluLeuLeuGluHisCysProSerGlnValThrValValAla 60

Db 121 AATGTGACTATTTGGGGTGGAGCTTCTGAAACACTGCGCTTCAACAGGTGACTGGAAGGCG 180  
 QY 61 GluLeuLeuLeuThrAlaSerAsnLeuThrValSerValLeuGluValGluGlyValPhe 80  
 Db 181 GAGCTGCTCAAGACAGATCAAACTCACTGTCTCTGTCTGGAGAGAGAGAGCTCTTT 240  
 QY 81 GluLeuGlySerPheLeuThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGlu 100  
 Db 241 GAAAGAGGCTCTTTAAGACACTTACTCTTCCATCACTACTCTGAAACAGTGCAGATGAG 300  
 QY 101 IleTyGluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThr 120  
 Db 301 ATTATGAGCTACGTGAACCGACGACGACGACGACGACGACGACGACGACGACGAC 360  
 QY 121 ArgLeuSerPheGluThrArgValArgIleSerValPheIleGlnThrAspValAlaLeuTyr 140  
 Db 361 CGCTTATCATTTAGACCAAGAGAAATATCTGTCTTCAATTCACAGACAGAGGCTTATAC 420  
 QY 141 LysProLysGlnGluValLysPheArgIleValThrLeuPheSerAspPheLysProTyr 160  
 Db 421 AAGCCAAAGCAAG 480  
 QY 161 LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGlnThrPhe 180  
 Db 481 AAAACCTCTTAAACATTTCTCATTAAGAGACCCCAAAATTCAAATTTGATCCAAACGTGTTG 540  
 QY 181 SerGlnGlnSerAspLeuGlyValIleSerValThrPheGluLeuSerSerHisPheIle 200  
 Db 541 TCACCAACAAAGTATCTTGAGCATTTCCAAACATTTTCAGTATCTTCCATCCAAATA 600  
 QY 201 LeuGlyAspTrpSerIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGln 220  
 Db 601 CTGGTGACCTGGCTCTTTCAGAGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660  
 QY 221 ValSerGluTyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSer 240  
 Db 661 GTTTCAGAAATATGATTAATCAAAATTTGAAAGTACCTTGACAGACACATTAATTTGTTCT 720  
 QY 241 MetAsnSerLeuHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProVal 260  
 Db 721 ATAAATTTCAAGCATTTAAATGATGATCATTACCGCAAGATATCAATATGGAGGCAAG 780  
 QY 261 LysGlyAspValThrLeuThrPheLeuProLeuSerPheTrpGlyLysLysValAsnIle 280  
 Db 781 AAGGAGACCTTACGCTTACATTTTATCTTTATCTTTGGGAGAGAGAGAGAGAGAGAT 840  
 QY 281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluGluMetLys 300  
 Db 841 ACAAAACATTTAAGATTAATGATGATCGCAAACTTCTTTTATGATGAGAGAGAGAA 900  
 QY 301 AsnValMetAsnSerSerSerAsnGlyLeuSerGluTyrLeuAspLeuSerProGlyPro 320  
 Db 901 AATGTATGATTTCTTCAATGAGATGATCTTCTGAAATACCTGATCTTATCTTCCCTGACCA 960  
 QY 321 ValGluIleLeuThrThrValThrGlnSerValThrGlyIleSerArgAsnValSerThr 340  
 Db 961 GTAGAAATTTTAAACACAGTGCAGAGATCAAGTTACGATTTTCAAGAAATGTAAGCACT 1020  
 QY 341 AsnValPhePheLysGlnHisAspTyrIleIleGluPhePheAspTyrThrThrValLeu 360  
 Db 1021 AATGTGTTCTTCAAGACACATGATTAATCATCATGATTTTATGATTAATCACTGCTTGG 1080  
 QY 361 LysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsnGlnLeu 380  
 Db 1081 AAGCCATCTCTCAACTTCAACGACGACGACGACGACGACGACGACGACGACGACGAC 1140  
 QY 381 ThrLeuGlnGluArgArgAsnValValIleThrValThrGlnArgAsnTyrThrGlu 400  
 Db 1141 ACTCTTGAG 1200  
 QY 401 TyrTrpSerGlySerAsnSerGlyAsnGlnLysMetGluValAlaGlnLysIleAsnTyr 420



Db 1201 TACTGAGCGGATCTAACAGTGAATCAGAAAATGAAAGCTGTTCAAGAAAATTAATAT 1260  
 QY ThrValProGlnSerGlyThrPheIysIleGluPheProIleuGluIuAspSerSerGlu 440  
 Db 1261 ACTGTCGCCCAAGTGAAGCTTTTAAGATTGAATTCCTCAATCTCGAGATTCAGTGA 1320  
 QY LeuGlnLeuIysAlaTyrPheLeuGlySerIysSerSerMetAlaValHisSerLeuPhe 460  
 Db 1321 CTACAGTTGAAGCTTAATTCCTTGGTAGTAAGAGTAGAGTGCAGTTCATAGTCTGTTT 1380  
 QY LysSerProSerIysThrTyrIleGlnLeuIysThrArgAspGluAsnIleIysValGly 480  
 Db 1381 AAGTCCCTAGTAAGCATACATCCAACTAAACAGAGATCAAAATATTAAGGTGGCA 1440  
 QY SerProPheGluLeuValIysSerGlyAsnIysArgLeuIleuGlnLeuSerTyrMetVal 500  
 Db 1441 TCGCTTTTGAAGTGGTGGTGAAGTGCACAAACGATGAGAGTGAATGATATGATGTA 1500  
 QY ValSerArgGlyGlnLeuValAlaValGlyIysGlnAsnSerThrMetPheSerLeuThr 520  
 Db 1501 GATATCCAGGGGACAGTTGGTGGCTAGGAAACAAATTCACAAATGTTCTCTTAAACA 1560  
 QY ProGluAsnSerTyrThrProIysAlaCysValIleValTyrTyrIleGluAspArgGly 540  
 Db 1561 CCAGAAAATTTCTTGAAGCTCCAAAGCCTGTGTAATGTGATTAATTAAGATGATGGG 1620  
 QY GluIleIleSerAspValLeuIysIleProValGlnLeuValPheIysAsnIysIleIys 560  
 Db 1621 GAATATTAAGTGAATGTCTTAATAATTCCTGTTCACTGTTTAAATAAGATGAAG 1680  
 QY LeuTyrTyrSerIysValIysAlaGluProSerGluIysValSerLeuArgIleSerVal 580  
 Db 1681 CTATATTTGAGTAAGTGAAGCTGAACCATCTGAGAAAGTCTCTTGAAGTCTGTG 1740  
 QY ThrGlnProAspSerIleValGlyIleValAlaValAspIysSerValAsnLeuMetAsn 600  
 Db 1741 AACAGCCCTGAGCTCCATAGTGGATTTGAGCTGTGACAAAGTGTGAATCTGATGAAAT 1800  
 QY AlaSerAspAspIleThrMetGluAsnValValHisGlnLeuGlnLeuTyrAsnThrGly 620  
 Db 1801 GCCTTAATGATATTAACAATGAGAAATGGTCCAGATGAGTGAACCTTATTAACACAGCA 1860  
 QY TyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTyrVal 640  
 Db 1861 TATTAATTTAGCATGTTCATGAATTTCTTTGCAAGTCTTTCAGGAAGTGCAGTGGGTA 1920  
 QY LeuThrAspAlaAsnLeuThrIysAspTyrIleAspGlyValTyrAspAsnAlaGluTyr 660  
 Db 1921 TTGACAGATGCAAACTCAGAAAGATTAATGATGATGTTATGACAAATGACAAATAT 1980  
 QY AlaGluArgPheMetGlnGluIuAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680  
 Db 1981 GCTGAAGGTTATGAGAGAAATGAGAGACATATTGTAAGATATTCATACATTTCTTTGG 2040  
 QY GlySerSerProHisValArgIysHisPheProGluThrTyrIleTyrLeuAspThrAsn 700  
 Db 2041 GGTAGAGTCCACATGTCGAAAGCATTTTCCAGACATTTGAGATTTGGCTAGACACCAAC 2100  
 QY MetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIleThrSerTyr 720  
 Db 2101 ATGGGTTACAGGATTTACCAAGAAATTTGAAGTAAGTACGTAAGTCTTCACTCTTGG 2160  
 QY ValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrThrProVal 740  
 Db 2161 GTGGCTACAGTGTGTTGTGATCTCTGAGACCTGGGCTTGAACCAACTCTCCACAGT 2220  
 QY GluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArg 760  
 Db 2221 GAGCTCCAGGCTTCCACACATTTTTCATTTTGAATCTTCCACTCTGTATACAA 2280  
 QY GlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuIysAspAlaThrGluVal 780  
 Db 2281 GGTGAAGAAATTTGCTTGAAGAAATTAATATATTTATTTGAAAGATGCCACTGAGTT 2340

QY LysValIleIleGluIysSerAspIysPheAspIleLeuMetThrSerSerGluIleAsn 800  
 Db 2341 AAGTATATCTGAGAAAGATGACAAATTTGATATTTCAATGACTTCAATGAATTAAT 2400  
 QY AlaThrGlyHisGlnGlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPhe 820  
 Db 2401 GCCACAGGCGACACAGACACCTTCTGTTCCCATGAGATGGGGCAACTGTTCTTTT 2460  
 QY ProIleArgProThrHisLeuGlyGluIleProIleThrValThrAlaLeuSerProThr 840  
 Db 2461 CCCATCAGGCCAACATCTGGAGAAATTCCTATCAGATCAGCTCTTTCACCCACT 2520  
 QY AlaSerAspAlaValThrGlnMetIleLeuValIysAlaGluGlyIleGluIysSerTyr 860  
 Db 2521 GCTTGTGATGCTGTGATCCCAAGATGATTTTAGTAAGGCTGAGAGAAATGAATAATCATAT 2580  
 QY SerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuIysThrLeu 880  
 Db 2581 TCACATCCATCTTATTAAGACTGATGACATAGCATAGATACCTGAAAACCTTTG 2640  
 QY SerPheSerPheProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIle 900  
 Db 2641 AGTTTCTGATTTCTCTCAATAACAGTGAAGTGCAGTGAAGAGTTCAGATCACTCAAT 2700  
 QY GlyAspValIleuGlyProSerIleAsnGluLeuAlaSerLeuIleArgMetProTyrGly 920  
 Db 2701 GGAGATGTTCTTGAGTCTTCCATCAATGAGCTTGAAGCTTCAATGATGGATGCTTATGGC 2760  
 QY CysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThr 940  
 Db 2761 TGTGTGAACAGCAATGATTAATTTGCTCCAAATATTTACATTTTGGATTATGACT 2820  
 QY LysIleIysGlnLeuThrAspAsnLeuIysGluIysAlaLeuSerPheMetArgGlnGly 960  
 Db 2821 AAAAAGAAACAACTGACAGATTAATTTGAAGAAAAGCTCTTTCATTAATGAGCAAGGT 2880  
 QY TyrGlnArgGluLeuLeuTyrGlnArgIuAspGlySerPheSerAlaPheGlyAsnTyr 980  
 Db 2881 TACCAAGAGAACTTCTTATCAGAGGAGAGTGGCTTTTCAAGTCTTTGGGAATAT 2940  
 QY AspProSerGlySerThrTyrLeuSerAlaPheValLeuArgCysPheLeuGluAlaAsp 1000  
 Db 2941 GACCTTCTGGAGACATTTGTTGTTGTTTGAATGTTTCTTGAACCCAT 3000  
 QY ProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTyrLeuIysGlyHis 1020  
 Db 3001 CCTTACATAGATATTGATCAGATGTGTAAACAGAACATACACTTGGCTTAAGACAT 3060  
 QY GlnIysSerAsnGlyGluPheTyrAspProGluArgValIleHisSerGlnLeuGlnGly 1040  
 Db 3061 CAGAAATCCACGGTGAATTTGGATCCAGAGAAAGATTCATAGAGACTTCMAAGT 3120  
 QY GlyAsnIysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArg 1060  
 Db 3121 GGCAATTAAGTCCAGTAACCTTACAGCTTATATTTGAGTCTCTCTGGATATAGA 3180  
 QY LysTyrGlnProAsnIleAspValGlnIleSerIleHisPheLeuGluSerIuPheSer 1080  
 Db 3181 AAGTATCAGCTTAACATTTGATGTCAGAGCTTATCACTTTTGGAGCTGAAATTCAGT 3240  
 QY ArgGlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGly 1100  
 Db 3241 AAGAGAAATTTCAAGAAATTAATATCTTACCTTAACTTAATGCAATTCATCACTGGG 3300  
 QY SerProIysAlaIysGluAlaLeuAsnMetLeuThrTyrPheAlaGluGlnGluIysGly 1120  
 Db 3301 AGTCTTAAGCAGAGAAAGCTTTGAATATGCTGACTTGAAGCAGCAAGAAAGTGGC 3360  
 QY MetGlnPheTyrValSerSerGluSerIysLeuSerAspSerTyrPheGlnProArgSerLeu 1140  
 Db 3361 ATGCAATCTGGGTGTATCAGAGATCCAAACTTTTCTGACTCTGGCAGACGCTCCCTG 3420

```

Qy 1141 AsplleGlValAlaAlaTyAlaleuSerHisphLeuGlnPheGlnThrSerGlu 1160
Db 3421 GATATGAAAGTTGACGCTTATGACTGCTCTCACACTTCTTACAAATTCACAGCTTGGAG 3480
Qy 1161 GlyIleProIleMetArgTrpLeuSerArgGlnArgGlnSerLeuGlyPheAlaSer 1180
Db 3481 GGAATCCCAATATGAGGTGGCTAAGCAGGCAAGAAATACCTGGGGTGGTTCATCT 3540
Qy 1181 ThrGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAlaAlaLeuMetAspThr 1200
Db 3541 ACTCAGGATACCACTGCTGCTTAAAGGCTCTGCTGTAATTTGACGCCCTTAATGAAATCA 3600
Qy 1201 GluArgThrAsnIleGlnValThrValThrGlyProSerSerProSerPro----- 1217
Db 3601 GAAAGGCAAAATATCCAAAGTACCGGTGACGGGGCTGAGCTCACCAAGCTCTGTAAGTTT 3660
Qy 1218 -----LeuAlaValAlaGlnPro 1223
Db 3661 CTGATTGACACACAAACCGCTTACTCTTACAGACAGAGCTTGTGGTACAGCCA 3720
Qy 1224 MetAlaValAlaSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValVal 1243
Db 3721 ATGGCAGTTAATATTTCCGCAAAATGTTTGGATTGCTATTGTCAGCTCAATGTTGTA 3780
Qy 1244 TyrAsnValLysAlaSerGlySerSerArgArgArgSerIleGlnAsnGlnLysAla 1263
Db 3781 TATATGATGGAAGGCTTCTGGGCTTCTTGAAGACGAGATCTATCCAAATCAGAAAGCC 3840
Qy 1264 PheAspLeuAspValAlaValLysGlnLysAsnLysAspLeuAsnHisValAspLeuAsn 1283
Db 3841 TTTGATTTAGATGTTGCTGTAAGAAATGAAATGATGATCCATCATCTGTCGATTTGAAAT 3900
Qy 1284 ValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluValAlaLeu 1303
Db 3901 GTGCTGACAACTTTTCCGGCCCGGGTACGAGTGCATGCTCTTATGAAATTAACCTTA 3960
Qy 1304 LeuSerGlyPheMetValProSerGluAlaIleSerLeuSerGluThrValLysVal 1323
Db 3961 TTAAGTGGCTTTATGATGCTCTTCAAGAACATTTCTTGACGAGACGATGAAAGAGTG 4020
Qy 1324 GluTyrAspHisGlyLysLeuAsnLeuTyrLeuAspSerValAlaGlnThrGlnPheCys 1343
Db 4021 GAATATGATCATGAGAAACCTCAACCTCTATTAAGATCTGTAATGAAACCACGCTTTTGT 4080
Qy 1344 ValAsnIleProAlaValAlaArgAsnPheLysValSerAsnThrGlnAspAlaSerValSer 1363
Db 4081 GTTAATATTTCTGCTGCTGAGAACTTTAAAGTTTCAAAATACCAAGATGCTTCACTGCTC 4140
Qy 1364 IleValAspTyrTyrGluProArgArgGlnAlaValArgSerTyrAsnSerGluValLys 1383
Db 4141 ATAGTGATTACTACTATGAGCCAAAGACAGCGCGGTGAGAACTTCAACTCTGAAAGTGAG 4200
Qy 1384 LeuSerSerCysAspLeuCysSerSerAspValGlnGlyCysAspProCysGluAspGlyAla 1403
Db 4201 CTGCTCCCTCTGACCTTGGACGATGTCAGAGGCTGCCCTCTTGAGAGATGAGAGCT 4260
Qy 1404 SerGlySerHisHisSerSerValIlePheIlePheCysPheLysLeuLeuTyrPhe 1423
Db 4261 TCAGGCTCCATCATCACTCTTCACTCACTTTATTTTCTGTTCAAGCTTCTGACTTT 4320
Qy 1424 MetGluLeuTrpLeu 1428
Db 4321 ATGGAACCTTGGCTG 4335

```

```

RESULT 2
AX534934 4761 bp DNA linear PAT 22-NOV-2002
LOCUS Sequence 1 from Patent WO02070696.
ACCESSION AX534934
VERSION AX534934.1 GI:25261474
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

```

```

REFERENCE
AUTHORS Schuh, A. and Sutherland, R.D.
TITLE Cation nucleic acid molecules polypeptides and methods of use
JOURNAL Patent: WO 02070696-1 12-SEP-2002;
Schuh, Andre (CA); Sutherland, Robert D. (CA)
FEATURES
source
1. .4761
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
113. .4450
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD57231.1"
/db_xref="GI:25261475"
/translators="MOGPPUTRAHLCTCTAALAVAPGPRPLVTPAGIIRREGANT
GVELLEHPCSQVTKALELKTASNLTVSVLEAGVERKDSFKTLTLPSPUNADBY
ELKAVGTQDEILFNSSTRSLSPETKRI SVFIQTDKALYRKQKVRIVTLFSDPKRY
KTSNLILKDPKSNLLOQMLSOQSDIGVSKTFOLSHSPILGMSIQOVNODQTYQS
FOSEYVLPKFEVTLQTPLYCSMNSKHLNGITFAKTYGKPVGVDVTLFPLSPFKG
KNITTFKINGSANSFNDDEMKVMDSNGISEYLDISSGPVILLTFTVESYVIG
SRNVSTNVPFKQHDYIIEFPDITVULKPSLNTFATYKTRADGNOULTERRNNVIT
VTQRNTERVWSGNSGNQKMEAVQKINVTVPQSGRIEFPILDESSLEQLKAYFLGS
KSSMAVHSLFKSPSKTYIOLKTRDENIKVSGPELVGSKRIKELSYWVSRGQVVA
VGKONSTMSLFPENSWTPKACIVYVYIEDDGEIISDYKIPOLVFKKIKLYMKV
KAEPSKVLRISSVTPDISIVGVADVCKVLMNANLIMENVHELELYNMGVIG
MPNNSFAYVQEGVLTLPANLTKVDYIDVYDNAAEAERFMENBEHVIDIHPSLGS
SPHVRKHPETWLTQDNGRIRIYQEFETVPDSTISWATGVISEDGLGLTTPV
ELQAFQPFILFNLPSYVIRGEFALFII FNLKQATVAKVIEKSDKFDLMTNE
INATGHOQTLVPSBEDATVLPFIRPETHGEIPIYVALSPDSADVTOMIILWAKGI
EKYSOSILDLIDNRLQSTLTKLSPSPNPTVGSERVOITRIGVLSPLASLI
IMPVSGGROMNIPAPNTIYLDYLLKKOOLDNTEKRLSPRQOYQBELYORBD
SRAPFNTPSSGTSALSAVLCFLEADYIDIDQVLRITTYLWLGKQKSNFDP
GRVHSELQGNKSPVTLVAVTSLGTRKQPNIDVQESHFLESSEFSRISDNT
LALITLALSVSQPKAKELNMTWABEQQQFQVNSSEKLSDSWQPSRLDI EYVA
YALSHLFOFTESEGIPIRMWLSRQNSLGSFASDTDTVALALALSEFALMNTERTN
IOVTVGSPSPVFKFLIDTHNRLLQTLAEVQPMANVISANGFPAICOLNVYV
VRSQSSRRRSRIONOEPDIDVAVKENDIDNHDNINVCSTSPSGRSGMALMEVNL
LSGFNVPSEALISLSETVKYRYEDHGLKLYLDSVNETQCVNIPAYNRKVSNTQAS
VSTVDYEPERQAVRSYNSEVKLSSCDLCSDVQCKRCEBDGASHHSSVIFPCPK
LTFMEMLM"
BASE COUNT 1420 a 912 c 1016 g 1413 t
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 4761
Score: 7326.50 Matches: 1427
Percent Similarity: 98.82% Conservative: 1
Best Local Similarity: 98.75% Mismatches: 0
Query Match: 99.71% Indels: 17
DB: 6 Gaps: 1
US-10-020-095-4 (1-1428) x AX534934 (1-4761)
Qy 1 MetGluGlyProProLeuLeuThrAlaAlaLeuLeuCysValCysThrAlaAlaLeu 20
Db 113 ATGAGGAGCCACCGCTCTGACCGCCGCCACTCTCTGCTGTCAGCCGCGCTG 172
Qy 21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGly 40
Db 173 GCCGTGGCTCCGGGCTGGTTCGTGTGACAGCCCGAGGATCATCAGCCCGAGGA 232
Qy 41 AsnValThrIleGlyValGluLeuLeuGlnHisCysProSerGlnValThrValLysAla 60
Db 233 AATGTGACTATTTGGGGTGGAGCTTCTGGAACAATGCTTCAAGAGTGACTGGAAGCG 292
Qy 61 GluLeuLeuLysThrLysAsnLeuLeuThrValSerValLeuGlnLysIleGluGlyValPhe 80
Db 293 GACGTCTCAAGACAGCATCAACCTCACTGCTGTCTCTGGAAGCAGAGAGGCTTTT 352
Qy 81 GluLysGlySerPheLysThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGlu 100

```

```
Db      353 GAAAAAGGCTCTTTAGACACTTACTCTTCATCTCACTCTCGAAGCAGTGAAGTAG 412
Qy      101 ILeTYGIuLeuAryValThrgIYArgThrGlnAspGluIleuPheSerAsnSerThr 120
Db      413 ATTATGAGCTACGTGTAAACCGAGCTTACCAGAGATGAGATTATATCTCTAATAGACC 472
Qy      121 ArgLeuSerPheGluThrIYArgIleSerValPheIleGlnThrAspLysAlaLeuYr 140
Db      473 CGCTATCATTTAGACCAAGAGAAATATGCTCTTCATTCAAACAGACAAAGCCTTATAC 532
Qy      141 LysProLysGlnGluValLysPheArgIleValIThrLeuPheSerAspPheLysProIYr 160
Db      533 AACCCAAAGCCAAAGAGTGAAGTTCCATGTTACCTCTTCACAGATTTTAAAGCCTTAC 592
Qy      161 LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGlnIleu 180
Db      593 AAAACCTCTTTAAACATTCATTAAGACCCCAAAATCAATTTGATCCAAACAGTGGTGG 652
Qy      181 SerGlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIle 200
Db      653 TCACAAACAAGTATCTTGAGTCAATTCCTCAAAACCTTTTCAGCTATCTCCCATCAATA 712
Qy      201 LeuGlyAspTyrSerIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGln 220
Db      713 CTGGGACTGGCTATTCATTCAGGTTCAAGTCAAGTCAAGACAGACATATTCATCATTTAC 772
Qy      221 ValSerGluTyrValLeuProLysPheGluValIThrLeuGlnThrProLeuTyrCysSer 240
Db      773 GTTTCGAATATGATATTAACCAAAATTTGAAGTATGATCTTGACACACCATTAATTTGTCT 832
Qy      241 MetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProVal 260
Db      833 ATGAATTCCTAAGCATTTAATAGTATACCATCAACGCAAGATATCATATGGGAAGCCAGTG 892
Qy      261 LysGlyAspValThrLeuThrPheLeuProLeuSerPheTyrGlyLysLysAsnIle 280
Db      893 AAAGGAGACGTAAACGCTTACATTTTACCTTTATCTTTGGGGAAGAAATAATATTT 952
Qy      281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluIleLys 300
Db      953 ACAAAACATTTAAGATTAATAGATGTGCAAACTTCTCTTTAATATGAGAGATGAAGA 1012
Qy      301 AsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSerProGlyPro 320
Db      1013 AATGTATATGATTTCTTCAATAGACCTTTCTGAATACCTGATTAATCTTCCCTGGACCA 1072
Qy      321 ValGluIleLeuThrThrValThrgIuSerValThrgIYIleSerArgAsnValSerThr 340
Db      1073 GTAGAAATTTTAAACACAGTGAAGATCAGTATACAGTATTTCAAGAAATGTAAACACT 1132
Qy      341 AsnValPhePheLysGlnHisAspTyrIleIleGluPhePheAspTyrThrThrValLeu 360
Db      1133 AATGTGTCTTCAAGACATGATTAATCATTCATTTGATTTGATTAATCTACTGTCTTG 1192
Qy      361 LysProSerLeuAsnPheThrAlaThrValLysValIThrArgAlaAspGluAsnGlnLeu 380
Db      1193 AAGCCATCTCTCAACTTCACAGCACTGTGAAGTAACTCGTCTATATGGCAACCAACTG 1252
Qy      381 ThrLeuGlnGluArgArgAsnAsnValValIleThrValThrgIuArgAsnTyrThrgIu 400
Db      1253 ACTCTTGAGAAAGAAAGAAATATATGATGATCATTAACGTGACACAGAAACTATATCTGAG 1312
Qy      401 TyrTrpSerGlySerAsnSerGlyLysAsnGlnLysMetGluAlaValGlnLysIleAsnTyr 420
Db      1313 TACTGAGGCGAGATCAAGTGAATCAAGAAATCAAGAAATGAGCTGTTTCAGAAATTAATTTAT 1372
Qy      421 ThrValProGlnSerGlyThrPheLysIleGluPheProIleuGlnLysAspSerSerGlu 440
Db      1373 ACTGTCCCAAGTGAAGCTTTTAAGATTTGAATTCCTCAATCTCGAGAGATTCCACTGAG 1432
Qy      441 LeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSerLeuPhe 460
Db      1433 CTACAGTTGAAGGCTATTTCTTGATAGTAAAGATGACATGCACTTACTGTGTT 1492
Qy      461 LysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLysValGly 480
Db      1493 AAGTCTCTAGTAAGACATATCAATCACTAAACAAAGAGAGAAATTAAGGGGGA 1552
Qy      481 SerProPheGluLeuValValSerGlyAsnLysArgLeuLysGluLeuSerTyrMetVal 500
Db      1553 TCGCCTTTGAGTGGTGGTGTAGTGGCAACAACATGGAAGATTAAGTATATGTA 1612
Qy      501 ValSerArgGlyGlnLeuValAlaValGlyLysGlnAsnSerThrMetPheSerLeuThr 520
Db      1613 GTATCCAGGAGGACAGTGTGGCTGAGGAAACAAATTCACAAATGTTCTCTTTAACA 1672
Qy      521 ProGluAsnSerTyrThrProLysAlaCysValIleValTyrTyrIleGluAspAspGly 540
Db      1673 CAGAAATTTCTTGACCTCCAAAGCGCTGTGAATGTGATTAATTAAGATGATGGG 1732
Qy      541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560
Db      1733 GAATTTATAGAGATGTTCTAATAAATCTCTGTTACGCTTTTAAATAAGATTAAG 1792
Qy      561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580
Db      1793 CTATATGAGATTAAGTGAAGTGAACCTGACCATCTGAGAAAGTCTCTTATAGATCTCTGG 1852
Qy      581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600
Db      1853 ACACAGCCTGACCTCAATGATGGAGTGTGAGTGTGACAAAGTGAATCTGATGAAT 1912
Qy      601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeuGluLeuTyrAsnThrGly 620
Db      1913 GCTCTAATGATATATTCATGGAATAATGTGGTCCATGAGTGGAACTTTATTAACACAGA 1972
Qy      621 TyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTyrVal 640
Db      1973 TATTAATTAAGCATGTTCATGAATCTTTTGACGCTTTCAGAAATGTGACTCGGGTA 2032
Qy      641 LeuThrAspAlaAsnLeuThrIYAspTyrIleAspGlyValTyrAspAsnAlaGluTyr 660
Db      2033 TTGACAGATGCAAACTCAACGAGATTAATATATGATGTGTTATGACAAATCACAATAT 2092
Qy      661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680
Db      2093 GCTGAGAGTTTATGAGAGAAATGAAGACATATGATATTCATGACTTTCTCTTG 2152
Qy      681 GlySerSerProHisValArgLysHisPheProGluThrTyrIleTyrLeuAspThrAsn 700
Db      2153 GGTACAGTCCACATGTCGAAAGCATTTTCCAGAGACTTGGAATTTGGCTAGACCAAC 2212
Qy      701 MetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIleHisSerTyr 720
Db      2213 ATGGGTTCACGATTTACCAAGAAATTTGAAGTACGTACGATTCATCATCACTCTCTGG 2272
Qy      721 ValAlaThrGlyPheValIleSerGluAspLeuGlyLysGluLysThrThrProVal 740
Db      2273 GTGGCTACTGGTTTGTGATCTCTGAGACCTGGGCTTTTGACATCAACATCACTCAAGT 2332
Qy      741 GluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArg 760
Db      2333 GAGCTCCAAAGCCTTCCAAACATTTTTCATTTTGTGAAATCTTCCCTACTCTGTATCAGA 2392
Qy      761 GlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrGluVal 780
Db      2393 GGTGAAGATTTGCTTGTGAATACTATATTCATATTTTGAAGAGATGCCCTGAGGTT 2452
Qy      781 LysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsn 800
Db      2453 AAGGTATCATTTGAGAAATGACAAATTTGATTTCAATGACTTCAAAAGAAATTAAT 2512
Qy      801 AlaThrGlyHisGlnGlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPhe 820
Db      2513 GCCACAGGCCACACAGACACCTTCTGTGTTCCAGTGAAGATGGGGCAACTGTTCTTTT 2572
```

Qy	821	ProilearProthrhialeuGlyGluileProileThrValileuSerProthr	840
Db	2573	CCCATCAGGCCAACATCTGGAGAAATTCCTATCAAGTCACAGGCTTTTCAACCACT	2632
Qy	841	AlaserapAlavalThrlMeclleuVallysalagluGlyileGluLyseTyr	860
Db	2633	GCTTCTGAATGCTGCACCCACATGATTTTAAAGGCTGAAGAAATGAATAATCATAT	2692
Qy	861	SerGlnSerileleuLeuAspLeuThrAspAsnArgleuGlnSerThrleuThrlleu	880
Db	2693	TCACATTCATCTTATTGACTTGACTGACATAGGCTACAGAGTACCCTTAAACCTTTG	2752
Qy	881	SerPheSerPheProProAsnThrValThrlGlySerGluArgValGlnileThrlAlle	900
Db	2753	AGTTTCATTTCCCTCAATATACAGTGAAGTGAAGTGAAGTTCAGATCATGCAATT	2812
Qy	901	GlyAspValleuGlyProSerilleuGlyleuAlaserleuileThrgMeCPotyGly	920
Db	2813	GGAGATGTTCTTGCTTCCTTCATCATAGGCTTAGGCTTCATGATTCGATCCCTTATGGC	2872
Qy	921	CysGlyGluGlnAsnMetileAsnPheAlaProAsnileTyrileuAspTyrleuThr	940
Db	2873	TGTGTGAACAGAACATATATAATTTGCTCAAAATATTACATTTTGATTTACTGACT	2932
Qy	941	LyshlysgGlnleuThrAspAsnleuLygGluysAlaserPheMetArgGlnGly	960
Db	2993	AAAAAGAAACAACGACAGATATATTTGAAAGAAAGCTTTTCATTTATAGGCAAGCT	2992
Qy	961	TyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyr	980
Db	2993	TACAGAGAGAACTTCTCTATCAGAGAGAGATGCTCTTCAGTGGCTTTTGGAATTTAT	3052
Qy	981	AspProSerGlySerThrTyrleuSerAlaPheValleuArgCysPheleuGluAlaAsp	1000
Db	3053	GACCTTCTGGAGACTTGCTGTGCAGCTTTGTTTAAAGTGTTCCTTGAAAGCCGAT	3112
Qy	1001	ProTyrileuAspPileuAspGlnAsnValleuHlsArgThrlTyrThrlTyrleuLygVhis	1020
Db	3113	CCTTACATAGATTTGATCAGATATGTGTTACAGAAATACATACCTGGCTTAAAGGACAT	3172
Qy	1021	GlnLyseSerAsnGlyGluPheTyrAspProGlyArgValileHisSerGlnleuGlnGly	1040
Db	3173	CAGAAATCAACGGTGAATTTTGGGATCCAGAAAGATGATTCATAGAGACTTCAAGGT	3232
Qy	1041	GlyAsnLyseSerProvalThrleuThrlAlaTyrileuAlaThrserleuLeuGlyTyrArg	1060
Db	3233	GGCAATAAAGTCCAGTACACTTACAGCTTATATGTAATTCCTCTCTGGATATATAGA	3292
Qy	1061	LystrGlnProAsnileuAspValGlnleuSerileHisPheleuGluSerGluPheSer	1080
Db	3293	AAAGTATCAGCTTACATTTGATGTCAGAGATCTATTCATTTTGGAGTTCGAATTCAGT	3352
Qy	1081	ArgGlylyleSerAspAsnTyrThrlleuAlaLeuileThrlTyrAlaLeuSerSerValGly	1100
Db	3353	AGAGCAATTTCAACAATTAATACCTTACCTTAACTTATGATGCTCATGCTGAGGG	3412
Qy	1101	SerProLyAlaLyGluAlaLeuAsnMetleuThrlTyrAspAlaGluGlnGluGly	1120
Db	3413	AGCTCTAAAGCGAAGAGCTTTGAATAGCTGAGTGAAGCGAAGCAAGAGAGTGGC	3472
Qy	1121	MetGlnPheTyrPvalSerSerGlySerLyseSerAspSerTyrGlnProArgSerleu	1140
Db	3473	ATGCAATTTCTGGGTCTCATGAGATCCAACTTTTGTGACTCTTGAGCCAGCTCCCTG	3532
Qy	1141	AspIleGluValAlaAlaTyrAlaLeuLeuSerHisPheleuGlnPheGlnThrSerGlu	1160
Db	3533	GATATTTGAAGTTGAGCTTATGACTGCTTCACACTTCTTACATTTTCAGACTTCTGAG	3592
Qy	1161	GlyIleProileMetArgTyrleuSerArgGlnArgAsnSerleuGlyGlyPheAlaSer	1180
Db	3593	GGAAATCCCAATTAATGAGGTGCTAAGCAGGCAAGAAATATGCTTGGGTGTTTGCATCT	3652
Qy	1181	ThrlGlnAspThrlThrlValAlaLeuLyAlaLeuSerGluPheAlaAlaLeuMetAsnThr	1200
Db	3653	ACTCAGATATCACTGGTGGCTTTAAAGGCTCTGTGAAATTTGCACCCCTTAATGAATACA	3712
Qy	1201	GluArgThrlAsnIleGlnValThrlValThrlGlyProSerSerProSerPro-----	1217
Db	3713	GAAAGCAAAATATCCAAATGACCGTGAACGGGGCTTACACAGTCTCTGTAAGTTT	3772
Qy	1218	-----LeuAlaValGlnPro	1223
Db	3773	CTGATTTGACACACAAACCGCTTACTCTTACAGACAGAGAGCTTGCTGTGTACAGCCA	3832
Qy	1224	MetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCysGlnleuAsnVal	1243
Db	3833	ATGCGAGTTAAATTTTCCGCAAAATGTTGGATTTGCTTATTTGCTCAGCTTCAATGTTGTA	3892
Qy	1244	TyrAsnVallysalaserGlySerSerArgArgArgSerIleGlnAsnGlnGluAla	1263
Db	3893	TATTAATGMAAGCTTCTGGGCTTCTTACAGAGAGATCTATCCAAATCAAGAGGCC	3952
Qy	1264	PheAspLeuAspValAlaVallysgIuAsnLyAspAspLeuAsnIleValAspLeuAsn	1283
Db	3953	TTTGATTTAGATGTTGCTGTGAAGAAATTAAGATGATCTCATATGATGATTTGAAT	4012
Qy	1284	ValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluValAsnLeu	1303
Db	4013	GTTGTATCAAGCTTTTTCGGGCCCGGGTAGAGATGCGATGCTTATGAAATTTAACTTA	4072
Qy	1304	LeuSerGlyPheMetValProSerGluAlaIleSerleuSerGluThrlVallysalVal	1323
Db	4073	TTAAGTGGCTTATATGCGCTTCAAGAGCAATTTCTTGACCAAGACAGTGAAGAAAGTG	4132
Qy	1324	GlyTyrAspIleGlyLyseleuAsnleuTyrleuAspSerValAsnGluThrlGlnPheCys	1343
Db	4133	GAATATGATATATGAAACCAACCTCTATTTGATTTCTGTAATATGAACCCAGTTTGT	4192
Qy	1344	ValAsnIleProAlaValAlaArgAsnPheLyValSerAsnThrlGlnAspAlaSerValSer	1363
Db	4193	GTTAATATTCCTCTGTGAGAAACTTTAAAGTTTCAATCCCAAGATGCTTCAGTGTCC	4252
Qy	1364	IleValAspTyrTyrGluProArgArgGlnAlaValArgSerTyrAsnSerGluValLyse	1383
Db	4253	ATAGTGATTTATATAGAGCCAGAGAGACAGCGGTAGAAATTTCAACCTTGAAGTGAAG	4312
Qy	1384	LeuSerSerCysAspLeuCysSerAspValGlnGlyCysArgProCysGluAspGlyAla	1403
Db	4313	CTGTCTCTCTGTATCTTTCAGTGAATGTCAGAGGCTCCGCTCTTGTAGATGAGAGCT	4372
Qy	1404	SerGlySerHisHisHisSerSerValIlePheIlePheCysPheLyseleuTyrPhe	1423
Db	4373	TCAGGCTCCCATCATCATCTTTCAGTCAATTTTATTTTCTGTTCAAGCTTCTGTACTTT	4432
Qy	1424	MetGluLeuThrlleu	1428
Db	4433	ATGGAATTTGGCTG	4447
RESULT 3			
AF410459			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Homo sapiens (human)			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE			
1. (bases 1 to 5883)			
AUTHORS			
Wu, X.F., and Schuch, A.C.			
TITLE			
Cell surface antigen CD109 is a novel member of the alpha(2)			
macroglubulin/C3, C4, C5 family of thioester-containing proteins			
JOURNAL			
Blood 99 (5), 1683-1691 (2002)			

MEDLINE 21849742  
 PUBMED 11861284  
 REFERENCE 2 (bases 1 to 5883)  
 AUTHORS Lin, M., Sutherland, D. R., Horsfall, W., Totty, N., Yeo, E., Nayar, R.,  
 TITLE Wu, X.-F. and Schubb, A. C.  
 JOURNAL Submitted (14-AUG-2001) Medicine, University of Toronto, 1 King's  
 College Circle, Room 7366, Toronto, Ontario M5S 1A8, Canada  
 FEATURES  
 source  
 1..5883  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="K1"  
 1..5883  
 /gene="CD109"  
 1..112  
 /gene="CD109"  
 113..4450  
 /gene="CD109"  
 /note="associated with the Goy alloantigen system"  
 /codon\_start=1  
 /product="CD109"  
 /protein\_id="AA184159.1"  
 /db\_xref="GI:19071209"  
 /translation="MQGPPLTAHLACVCTAALAVAPGRPLVTAPGIIRGNVTI  
 GVLEHCPQVTVKAEILKTSNLTIVSLAEVFEKGSFKTLPLSPINSADEIY  
 ELRTVGRTODEILFNSSTRLSPEKRSIVFIOTDKALYKQEVKRIYTLPSDKPY  
 KSNITLIDPKSNLLOWLSSQSDLGIVSKTPOLSSHPILIDMSIQVQNDQYQS  
 ROVSRYLPEFVYTIOTPLVCSMSKHLNGITTAKTGYKPVKGDITFLPLSTMG  
 KNAITTEFKINGSNDYIEPFDDYTVLKPSLNTFAVYKLVRAAGNOLTEERNNVIT  
 VQNRVTEYWSGNSGNQKMAVOKINTVPOSGTFKIEPILDESSELQKAYFLGS  
 KSMVAHSEFKSPSKTYIOLKTRDENIKVSGPELVSGNKKLSLWVVRGOLVA  
 VKQKSTWELTPENSWTPKACVYIYIEDDEIISDVKITVQYVFKKKIKLWYSKY  
 KAPSEKVSRLISVTOPSIVGIVAVDKSVNMAANDITMENVAHELELVNTGYLG  
 MENSFAVQECGLWLDANLTKYIDIVYDVAEAEEMENSGHIVDHDPLSGS  
 SHAPKHPPEITWIDTMNGYRIQEFETVPDSITSWATGVISEDGLATTPV  
 ELQAPQPFPIPLNPSYVGRSEPALEITIPYVLDKATVKKVIEKSDKFDILMTSE  
 INATGHOOTLIVPSBDGATVLPFRPTHGELPIYVTLASPTASDAVVMILVKAEG  
 EKSYSOGLIDITDNRLQSTLKTLSFPFNPVTGSEKQITAGIDVQAPSLNGIASL  
 IIMPYGCGRQMINFAPNIVLIDYLTKKQLTDNLKELKSPRQGYQRELLYQEDG  
 SFSAFGNVDPSGTSVAFVLCPLFLEADYIDIDQNVLRITTYMLKQKSGEPMW  
 GVIHSELDQGNKSPVTLTAIVTSLGKRYKOPNVQGESIHPLESESRGISNNT  
 IALITYLASSVGSPPAKKALNMLTWRAEGGOMFVSESLSDSMOBRSLDIVAA  
 VALLSHLOFQPSSEGIPIRMWLSRQNSLGGFASODTTVAALKASERPALNTERVN  
 IOVTVTGSSPSPVFKLIDITHRLQLTRELAVOPMAVNISANGGFALCOLNVVN  
 VASGSSRRRSIQOEAFDLVAVKENDDLNVDLVNCTSFSGRSGNALMEVNL  
 LSGFWPSEAIISLSTVKVEYDHGKLNLYLDSVNETQCVNI PAVRNKVSNTODAS  
 VSIVDYVEPRQAVRSYNSSEVKLSSCDLCSVDVQGRPCEDGASGSHHSVIFIFCFK  
 LLYPMELMLT"  
 4451..5883  
 /gene="CD109"  
 4748  
 /note="alternative polyA site found on clone K1"  
 /gene="CD109"  
 BASE COUNT 1697 a 1173 c 1241 g 1772 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 0 Length: 5883  
 Score: 7326.50 Matches: 1427  
 Percent Similarity: 98.82% Conservatives: 1  
 Best Local Similarity: 98.75% Mismatches: 0  
 Query Match: 99.71% Indels: 17  
 DB: 9 Gaps: 1  
 US-10-020-095-4 (1-1428) x AF410459 (1-5883)  
 QY 1 MetGInGlyProProleuLeuThraIaahIbLeuLeuCyVaIaCyThraIaahIaLeu 20  
 Db 113 ATGCAAGGCCACCGCTCTGACCGGCCACCTCTCTGCGTGGACCGCCGCGCTG 172

QY 21 AlaValAlaProGlyProArgPheLeuValThraIaProGlyIleIleArgProGlyGly 40  
 Db 173 GCCGTGGCTCCCGGGGCTCGGTTTCTGGTGCACGCCAGGAGATCATCAGGCCGAGAGA 232  
 QY 41 AsnValThrIleGlyValGluLeuLeuLuhIscYProSerGlnValThraValAla 60  
 Db 233 AATGAGCTATTGGGGGTGAGCTCTGTGAAACATGCTCTTACAGTACTGTGAAAGCG 292  
 QY 61 GluLeuLeuYsThraIaSerAsnLeuThraValSerValIleuGluAlaGlyValPhe 80  
 Db 233 GAGCTGCTCAAGACAGCATCAAACTCATCTGTCTGTCTGGAAGCAGAAAGATCTTT 352  
 QY 81 GlyIysGlySerPheYsThreIuThreProSerIleuProLeuAsnSerAlaaspGlu 100  
 Db 353 GAAAAAGCTCTTTAAGACACTTACTTCCATCATCACTCTGACAGTGCAGATGAG 412  
 QY 101 IleYrGluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThr 120  
 Db 413 ATTATGAGCTACGCTGTAACCGGAGTACCCAGAGTGAATTTATCTCTAATGTAC 472  
 QY 121 ArgLeuSerPheGluThrIleYsArgIleSerValPheIleGlnThrAspIleValLeuYr 140  
 Db 473 CGCTTATCATTTGAGACCAAGACGAAATATCTGTTCTTCAATCAACAGACAGGCTTATAC 532  
 QY 141 LysProIysGlnGluValIleYsPheArgIleValThreIuPheSerAspPheYsProYr 160  
 Db 533 AAGCAAAAGCAAGAAAGTAAAGTTTGCATTTGACCTTCTGATTTTAAGCTTAC 592  
 QY 161 LysThrSerLeuAsnIleLeuIleYsAspProIysSerAsnLeuIleGlnGlnThrLeu 180  
 Db 593 AAACCTCTTAAACATTCATTAAGAACCCCAATCAAAATTTGATCAACAGGCTTG 652  
 QY 181 SerGlnIserAspLeuGlyValIleSerYsThrPheGlnLeuSerSerHISproIle 200  
 Db 653 TCACACAAAGTATGATCTTGAGTCAATTCACAAATTTTCAGTATCTTCCATCACTAATA 712  
 QY 201 LeuGlyAspIlePheIleGlnValGlnValAsnAspGlnThrYrYrGlnSerPheGln 220  
 Db 713 CTGGTGACTGGTCAATTCAGTTCAAGTGAATGACGACATNTATCATCATTTTCAG 772  
 QY 221 ValSerGluYrValLeuProIysPheGluValThreIuGlnThrProLeuYrCysSer 240  
 Db 773 GTTTCAGATATGATTTACCAAAATTTGAAGTGACTTGCAGACACCATTAATATTGCT 832  
 QY 241 MetAsnSerIysHisLeuAsnGlyThrIleThraIaYsYrThrYrGlyYsProVal 260  
 Db 833 ATGAATTTTAAACATTTAAATGATACATCAACGGAAGTATACATATGGAAGCAGTG 892  
 QY 261 LysGlyAspValThrLeuThrPheLeuProLeuSerPheTProGlyYsLeuYsAsnIle 280  
 Db 893 AAAGAGACGTAACGCTTACATTTTACCTTATCTTTGGGAAAGAAAGAAAATATT 952  
 QY 281 ThrIysThrPheYsIleAsnGlySerAlaAsnPheSerPheAsnAspGluIleuYs 300  
 Db 953 ACAAACATTTAAGATTAATGATCTGCAAACTCTCTTTAATGATGAGAGATGAAA 1012  
 QY 301 AsnValMetAspSerSerAsnGlyLeuSerGluYrIleuAspLeuSerSerProGlyPro 320  
 Db 1013 AATGTAATGATTTCTTCAATGAGCTTCTCAAAATCCGGAATCATCTTCCCTGGACCA 1072  
 QY 321 ValGluIleLeuThrThraValThrcIuSerValThrgIyIleSerArgAsnValSerThr 340  
 Db 1073 GTAGAAATTTTAAACACGATACAGATCAAGTACAGGATATTTCAAGAAATGTAAGCACT 1132  
 QY 341 AsnValPhePheYsGlnHisAspYrIleIleGluPhePheAspYrThrThraValLeu 360  
 Db 1133 AATGTTCTTCAAGCAATGATTAATCATGAGTTTGTGATTAATACATGCTCTTG 1192  
 QY 361 LysProSerLeuAsnPheThraIaThrValYsValThraGluIleAspGlyAsnGlnLeu 380  
 Db 1193 AAGCATCTTCAACTTCAAGCAGCACTGTGAAGTAACTCTGCTGATGAGCAACCAACTG 1252  
 QY 381 ThrLeuGluGluArgArgAsnValValIleThrValThrgIuArgAsnYrThrGlu 400

D	1253	ACTCTTGAAGAAAGAAATATGTACTATACAGACACAGAGAACTATACAG	1312	D	2333	GAGCTCCAGCCTTCCAAACCATTTTTCATTTTGTGAATCTCCCTACTGTATACAGA	2392
Q	401	TYRTPSERGlySerAsnSerGlyAsnGlnlyMetGluAlaValGlnlyLeuAsnTyr	420	Q	761	GlyGlnGlnPheAlaLeuGlnLeuIleThrIlePheAsnTyrLeuIleAspAlaThrGluVal	780
D	1313	TACTGACGGAGTCTAACAGCTGGAATACAGAAATGGAAGCTGTTCCAGAAATTAATAT	1372	D	2393	GGTGAAGAAATTTGCTTTGGAAATACATATATTCATATTTTGAAGAATGCCATGAGTT	2452
Q	421	ThrValProGlnSerGlyThrPheIleGlnPheProIleLeuGlnAspSerGln	440	Q	781	LysValIleIleGlnLysSerAspLysPheAspIleLeuMetThrSerSerGlnIleAsn	800
D	1373	ACTGCCCCCAAGTGAACCTTTTAAGATGTAATCCCAATCCCTGAGATTTCCAGTGAG	1432	D	2453	AAGGTATCATTTGAGAAAGTGAACAAATTTGATATTTCTAATGACTTCAATGAAATGAAT	2512
Q	441	LeuGlnLeuValaIleTyrPheLeuGlySerIleSerSerMetAlaValHisSerLeuPhe	460	Q	801	AlaThrGlyIleGlnGlnThrLeuLeuValProSerGlnAspGlyAlaThrValIleuPhe	820
D	1433	CTACAGTGAAGCCCTATATTTCTTGGTAGTAAAGTACATGACATGACATAGCTGTGTT	1492	D	2513	GCCACAGGCCACACAGACCTTCTGTGTTCCAGTGAAGATGGGCACTGTTCTTTT	2572
Q	461	LysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLysValGly	480	Q	821	ProIleArgProThrIleLeuGlyGlnIleProIleThrValThrAlaLeuSerProThr	840
D	1493	AAGTCTCTAGTAAACATACATCTCCAACTAAACACAGATGAAATATTAAGGTGGGA	1552	D	2573	CCCATGAGCCACACATCTGGGAGAAATTCATACAGTCAAGTCTTTACACCAC	2632
Q	481	SerProPheGlnLeuValValSerGlyAsnLysArgLeuLysGlnLeuSerTyrMetVal	500	Q	841	AlaSerAspAlaValThrGlnMetIleLeuValLysAlaGlnGlyIleGlnLysSerTyr	860
D	1553	TCGCTTTTGAAGTGTGTAGTGCAACAAGATTGAAGATTGAAGCTATATGTGTA	1612	D	2633	GCTTGTATGCTGTCAACCCAGATGATTTTGTAAAGGCTGAGAGAAATGACATAT	2692
Q	501	ValSerArgGlyGlnLeuValAlaValGlyLysGlnAsnSerThrMetPheSerLeuThr	520	Q	861	SerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeu	880
D	1613	GTATCCAGGGGACAGTGTGTGTGTAGGAAACAAATTCACAAATGTTCTCTTTAACA	1672	D	2693	TCACAATCCATCTTATTAAGCTTGACTGACAAATAGCTACAGAGTACCTGAAACCTTG	2752
Q	521	ProGlnAsnSerTyrThrProLysAlaCysValIleValTyrTyrIleGlnAspAspGly	540	Q	881	SerPheSerPheProProAsnThrValThrGlySerGlnArgValGlnIleThrAlaIle	900
D	1673	CCAGAAATTTCTTGACCTCCAAACCTGTGTATGTGTATGTATATTGAAGAGATGGG	1732	D	2753	AGTTTCATTTCTCTCTATATACATGACCTGGCAGTGAAGATTCAGATCACTGCAT	2812
Q	541	GlnIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys	560	Q	901	GlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGly	920
D	1733	GAAATATTAAGATGTTCTTAAATTCCTGTTCAGCTGTTTAAATAAATTAAGTAAAG	1792	D	2813	GGAGATGTTCTGTGCTTCCATCAATAGCTTACCTCATTTGATGGATGGCTTATGCG	2872
Q	561	LeuTyrTyrPserLysValLysAlaGlnProSerGlnLysValSerLeuArgIleSerVal	580	Q	921	CysGlyGlnGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThr	940
D	1793	CTATATTTGAGTAAAGTGAAGCTGAAACATCTGAGAAAGCTCTCTTGAAGATCTGTG	1852	D	2873	TGTGTGAACAGAACATGAATTAATTTGTCTCCAAATTAATTAATTTGATTAATGACT	2932
Q	581	ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn	600	Q	941	LysLysLysGlnLeuThrAspAsnLeuLysGlnLysAlaLeuSerPheMetArgGlnGly	960
D	1853	ACACAGCTGACCTCATAGTGTGGATGTAGCTGTGCAAAAGTGAATCTGATGAT	1912	D	2933	AAAAAGAAACATGACAGATTAATTTGAAGAAAGAAACCTTTTCATTTATGAGCAAGT	2992
Q	601	AlaSerAsnAspIleThrMetGlnAsnValValHisGlnLeuGlnLeuTyrAsnThrGly	620	Q	961	TyrGlnArgGlnLeuLeuTyrGlnArgGlnAspGlySerPheSerAlaPheGlyAsnTyr	980
D	1913	GCCTCTATGATATTAACAATGGAATATGTGCTCCATGAGTTGGAACCTTTAACAAGCA	1972	D	2993	TACAGAGAACATCTCTCATACAGAGGAAGATGGCTTTTCAGTGTGTTGGAATAT	3052
Q	621	TyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGlnCysGlyLeuTyrVal	640	Q	981	AspProSerGlySerThrTyrPleuSerAlaPheValLeuArgCysPheLeuGlnAlaAsp	1000
D	1973	TATATTTAGGCATGTTCATGAATCTTTTCAGACTTTTCAGAAATGTGACTGTGGTA	2032	D	3053	GACCTTCTGGAGCACTGTGTGTGACGCTTTGTTTGAAGATTTCTTGAAGCCGAT	3112
Q	641	LeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyrAspAsnAlaGlnTyr	660	Q	1001	ProTyrIleAspIleAspGlnAsnValIleuHisArgThrTyrThrTyrPleuLysGlyHis	1020
D	2033	TTGACAGATGAAACCTCACAGAGATTAATTAATGATGTGTTTGAACAAATGAGAAAT	2092	D	3113	CTTTACATTAATATTAATGATGCAATGTGTTACACAGAAATACATTCGCTTAAGACAT	3172
Q	661	AlaGlnArgPheMetGlnGlnLysGlnGlyHisIleValAspIleHisAspPheSerLeu	680	Q	1021	GlnLysSerAsnGlyGlnPheTyrAspProGlyArgValIleHisSerGlnLeuGlnGly	1040
D	2093	GCTGAGAGTATTATGAGAGAAATGAAGACATATTAATGATCAATGCAATTTCTTTG	2152	D	3173	CAGAAATCCACAGGTGATTTGGATCCAGAAAGTATGATCATAGTAGCTTCAAGGT	3232
Q	681	GlySerSerProHisValArgLysHisPheProGlnThrTyrIleTyrPleuAspThrAsn	700	Q	1041	GlyAsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArg	1060
D	2153	GGTAGCAGTCCACATGTCCGAAAGCATTTTCAGAGACTTGGAATTTGGCTAGACACAC	2212	D	3233	GGCAATTAAGATCCAGTACATTAACAGCTTAATTTGTAATTTCTTCCGGAGTATAGA	3292
Q	701	MetGlyTyrArgIleTyrGlnGlnPheGlnValThrValProAspSerIleThrSerTyr	720	Q	1061	LysTyrGlnProAsnIleAspValGlnGlnLysSerIleHisPheLeuGlnSerGlnPheSer	1080
D	2213	ATGGGTAAACAGATTTTACAAAGAAATTTGAAGTAACTGATCTATATCTTCTTGG	2272	D	3293	AAGTATACGCTTAACATTAATGATGCAAGAGTCAATTCATTTTGGAGTGTGAATTCAG	3352
Q	721	ValAlaThrGlyPheValIleSerGlnAspLeuGlyLeuGlyLeuThrThrProVal	740	Q	1081	ArgGlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGly	1100
D	2273	GTGGCTACTGTTTGTGATCTCTGAGAGACCTGGGTCTTGAGACTAACAATCTCCAGTG	2332	D	3353	AGAAGATTTCAACACATTAATTAATCTAGCCCTTAATTAATGATGTCATCAGTGGG	3412
Q	741	GlnLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArg	760	Q	1101	SerProLysAlaLysGlnAlaLeuAsnMetLeuThrTyrArgAlaGlnGlnGlnGly	1120
				D	3413	AGTCTTAAGCGAAGAGCTTGAATATGCTGACTTGGAGACAGAACAAAGAGTGGC	3472

QY 1121 MetGlnPheTyrValSerSerGlnSerIleuSerAspSerTyrGlnProArgSerIleu 1140  
 DB 3472 ATGCATTCCTGGGTCATCATGAGTCACAACTTTCTGACTCTGGCAGCAGCTCCCTG 3532  
 QY 1141 AspIleGluValAlaAlaTyrAlaIleuSerHisPheLeuGlnPheGlnThrSerGlu 1160  
 DB 3533 GATATTTGAAGTTGACGGCTATGACCTGCTCACAACCTTTTCAATTTCAGACTTCTGAG 3592  
 QY 1161 GlyIleProIleMetArgTyrIleuSerArgIleuArgSerLeuGlyIlePheAlaSer 1180  
 DB 3593 GGAATTCCTCAATTATGAGTGGCTTACAGGCAAGAAATAGCTTGGGTTGCTTTGCTAT 3652  
 QY 1181 ThrGlnAspThrThrValAlaIleuIleuSerGlnPheAlaAlaIleuMetAsnThr 1200  
 DB 3653 ACTCAGATATACACTGTGCTTTAAAGCTGCTGTGCTGATTTTGACGCTTATGATATA 3712  
 QY 1201 GluArgThrAsnIleGlnValThrValThrGlyProSerSerProSerPro----- 1217  
 DB 3713 GAAAGGACAAATATCCAGTGAACGGGCGCTGACCTCAACAACTCTGTAAAGTTT 3772  
 QY 1218 -----LeuAlaValAlaGlnPro 1223  
 DB 3773 CTGATTGACACACACCGCTTACTCTTCAAGACAGACCTTGTGTGATACGCCA 3832  
 QY 1224 MetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCysGlnIleuAsnValVal 1243  
 DB 3833 ATGGCAGTTAATATTTCCGCAATGGTTTGGATTGGCTATTTGCTAGCTCAATGTGTA 3892  
 QY 1244 TyrAsnValIleAlaSerGlySerSerArgArgArgArgSerIleGlnAsnGlnGluAla 1263  
 DB 3893 TATAATGTCAAGGCTTCTGGGCTCTTCAAGACAGCAAGATCTATCAAAATCAAGAGCC 3952  
 QY 1264 PheAspIleuAspValAlaValIleGluIleuIleuAspIleuAsnIleValIleuIleu 1283  
 DB 3953 TTGATTATGATGCTGCTGTAAAGAAATTAAGATCTCAATCATGTGATTTGAT 4012  
 QY 1284 ValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaIleuMetGluValAsnIleu 1303  
 DB 4013 GTGTGACAAAGCTTTTCGGGCCCGGGTNGAGATGGCATGGCTTTATGAGATTAACTTA 4072  
 QY 1304 LeuSerGlyPheMetValProSerGluAlaIleSerIleuSerGluThrValIleuVal 1323  
 DB 4073 TTAAATGGCTTATGTGCTCTTCAAGACAAATTTCTGAGCAGACAGTGAAGAAAGTG 4132  
 QY 1324 GluTyrAspHisGlyIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 1343  
 DB 4133 GAATATGATGATGAGAAACTCAACCTCTATTTAGATTCTGTAATGAAACCCAGTTTGT 4192  
 QY 1344 ValAsnIleProAlaValArgAsnIleuValSerAsnThrGlnAspIleuValSer 1363  
 DB 4193 GTTAATATTTCTGCTGTGAGAACTTTAAAGTTTCAATATCCAAAGATCTTCAAGTTC 4252  
 QY 1364 IleValAspTyrTyrGluProArgArgGlnAlaValArgSerTyrAsnSerGluValIle 1383  
 DB 4253 ATAGTGTGATTAATATGAGCCAAAGAGACAGCGGTGAGAAAGTTTACACTGAAAGTAA 4312  
 QY 1384 LeuSerSerCysAspIleuIleuSerAspValGlnGlyCysArgProCysGluAspGlyAla 1403  
 DB 4313 CTGTCTCTCTGAGCTTTGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 4372  
 QY 1404 SerGlySerHisHisIleSerSerValIlePheIlePheCysPheIleuIleuTyrPhe 1423  
 DB 4373 TTAGGCTCCCATCATCATCTTCACTGATTTTATTTTCTGTTTCAAGCTTCTGATCTT 4432  
 QY 1424 MetGluLeuTyrIleu 1428  
 DB 4433 ATGGAACCTTGGCTG 4447  
 RESULT 4  
 AX534938  
 LOCUS AX534938 5895 bp DNA linear PAT 22-NOV-2002  
 DEFINITION Sequence 5 from Patent WO02070696.

ACCESSION AX534938  
 VERSION AX534938.1 GI:25261482  
 KEYWORDS  
 SOURCE  
 ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE  
 1 Schuh, A. and Sutherland, R. D.  
 Cdi09 nucleic acid molecules, polypeptides and methods of use  
 Patent: WO 02070696-A 5 12-SEP-2002;  
 Schuhn, Andre (CA); Sutherland, Robert D. (CA)  
 location/Qualifiers  
 1..5895  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 113..4450  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAD57233.1"  
 /db\_xref="GI:25261483"  
 /translation="MQGPELITAAHLICVTAALAVAPRPFLVTAGIIRPGNVTI  
 GVELLHRCPSQVTVKAEILKTASMLTVSVLEAEVFEKSEFKTLTLPGLPLNSADEIY  
 ELRYGRQDEILFNSNTRLSFEKRLSVFLOTKALYKPOEVFRIVTLFSDPKPY  
 KTSNLILIKDKNMLIQOMLSQSDGLVISTKPOLSHPLIGDMSIQVNDQFYOS  
 FOSEVYLPKFEVTLQPLVCSNMSKHLNGITAKTKYGRKPKGVTLLPLPLSMFK  
 KKNITKPKINGSANFSNDEEMQVNDSSGLSEYDLSPGPEIITTYTESVTGI  
 SRNVTNFFKQHYTILEFDYTVLVRKSLFVATVKTADGQQLIEERNVAVIT  
 VTQNVTEYWGSSNGQKMAVOKINATVLEQSGTFKLEPILDDSSLOKAVPLGS  
 KSMVAHSLFKSPSKTYLOLKTREBNIKVSGFELVSGNRKLEISPMVSRQLVA  
 VGNKSNMFSLTLPNSMTPKACVIVYLEDGEILSDYLKLPOLVFENKLIKLYMSK  
 KAPSEKYSLSISVTPQPSIVGIVAVKSVNLVMAANDITMENVHELEIYNTYYIG  
 MNMSEAFVQSCGLMTLDATDXYDIDGVADNAEYERFMENEGHVIDHDSLGS  
 SPVRKPEPBTWLDITMGRTITQEFVTPDSITSVWAGFVISEDLGLITTPV  
 ELQAPQPTFLNLPYSVIRGEERALETTINYLKDAIEVYVILEKSDKPIIMTSE  
 INATQOQTLVPSDEGATVLPPIRPTHGSIPIVTALESPADAVTOMLVAGGI  
 EKSYSOGLDLTQNRLOSTLKTLSFSPPTVGSERVQITAGDVGSPINGLAL  
 IRMPYSGEONMINPAPNIVYLDVLTFRKQTLNDLKEKALSFMQGRRELLYREDG  
 SPSAFNGYDPSGSTMASFVLRCLPADPIDIDQNTLHRTYTLKHKQKNGEMDP  
 GRVHSELQGNKSPVTLTATVTSILGKRYQPNIDVESIHLELESFGISDNT  
 LALITVALSVSGSPKAKELMLTWAEQEGMQFVWSSEKRLDSQPRSLDI EVA  
 YALISHLPQFQTSBGIPIMRWLSRQNSLQSGASTODTVALKLSSEPAALMNTERTN  
 IOVTVTPGSSSPKPLIDTNRLRLLOTAELAVOPAVNVSANGFGALICOLVVYN  
 VKAGSSRRRSIONOARFDLDVAVKKNLVDNVCNCSBSPGSGMALEVN  
 LSGMVPSEBALSSETYAKVYDQKLNLYLDSVNETQFCVNTAVNRFKISNTODAS  
 VSIQVYEPERQAVRSYNSVSKSSCDLCSDVQGRPCEDASGSHHSVIFLFCFK  
 LLTYFMEMLL"  
 BASE COUNT 1709 a 1173 c 1241 g 1772 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 0 Length: 5895  
 Score: 7326.50 Matches: 1427  
 Percent Similarity: 98.82% Conservative: 1  
 Best Local Similarity: 98.75% Mismatches: 0  
 Query Match: 99.71% Indels: 17  
 DB: 6 Gaps: 1  
 US-10-020-095-4 (1-1428) x AX534938 (1-5895)  
 QY 1 MetGlnGlyProProIleuLeuThrAlaAlaHisIleuIleuCysValCysThrAlaAlaIleu 20  
 DB 113 ATGCAGGGCCCAACCGCTCTTACCGCCCAACCTCTCTGCTGTGACAGCGCGGCTG 172  
 QY 21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGlyGly 40  
 DB 173 GCCGTGGCTCCCGGGCTCGGTTTCTGTGTGACAGGCCCAAGGATATCATGAGCCGGA 232  
 QY 41 AsnValThrIleGlyValGluLeuLeuGluHisCysProSerGlnValThrValIleVal 60  
 DB 233 AATGTGACTAATTTGGGCTGAGACTTCTGTGACACATGCGCTTCAAGTGTGAGAGCG 292



QY 61 GluLeuLeuYsThrAlaSerAsnLeuThrValIleuGluValIleGluValIlePhe 80  
 Db 293 GAGTCGCTCAAGACAGCATCAAACTCTCTCTGCTCGAAGACGAAAGAGTCTTT 352  
 QY 81 GluIysGlySerPheIysThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGlu 100  
 Db 353 GAAAGAGGCTCTTTAAGACACTTACTCTTCATCCTCACTGAAACAGCGAGTAC 412  
 QY 101 IleTyrGluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnThr 120  
 Db 413 ATTATAGCTACGTGTACCGGACGTACCCAGATGAGATTATTTCTCTAATAGTACC 472  
 QY 121 ArgLeuSerPheGluThrIysArgIleSerValPheIleGlnThrAspValIleLeuTyr 140  
 Db 473 CGCTTATCATTTGACCAAGAGATATCTCTCTTCAATCAACAGACGAGCTTATAC 532  
 QY 141 LysProIysGlnIleValIysPheArgIleValIleThrLeuPheSerAspPheIysProTyr 160  
 Db 533 AAGCCAAAGCAAGAGTAAAGTTTCGCAATGTACACTCTTCAATTTTAAACCTTAC 592  
 QY 161 LysThrSerLeuAsnIleLeuIleLysAspProIysSerAsnLeuIleGlnGlnTyrLeu 180  
 Db 593 AAAACCTCTTAAACATTTCTCATTAAGACCCCAATCAATTTGATCCAAAGTGGTTG 652  
 QY 181 SerGlnIysSerAspLeuGlyValIleSerIysThrPheGlnLeuSerSerHisProIle 200  
 Db 653 TCACAAACAAAGTGCATTTGAGTCATTTCCAAAATTTTCAGTATCTTCCATCCATA 712  
 QY 201 LeuGlyAspTyrPheIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGln 220  
 Db 713 CTGGTGTACGTGTATTCAGATTCAGTGAATGATACAGACATTTATCATCATTTTCAG 772  
 QY 221 ValSerGluTyrValIleProIysPheGluValIleThrLeuGlnThrProLeuTyrCysSer 240  
 Db 773 GTTTCAGAAATGTGATTAACCAAAATTTGAAAGTGCATTTGCAGACACCATTAATGTTCT 832  
 QY 241 MetAsnSerIysHisIleAsnGlyThrIleThrAlaIysTyrThrTyrGlyLysProVal 260  
 Db 833 ATGAAATCTAAGCATTTAATATGTATACCATCAAGCAAAAGTATCATATGGGAAGCCAGTG 892  
 QY 261 LysGlyAspValIleThrLeuThrPheLeuProLeuSerPheTyrGlyLysValAsnIle 280  
 Db 893 AAAGAGACGTAAAGCTTAACTTTTACCTTTATCTTTGGGAGAAAGAAATAATTT 952  
 QY 281 ThrIysThrPheIysIleAsnGlySerAlaAsnPheSerPheAsnAspGluGluMetLys 300  
 Db 953 ACAAAACATTTAAGATTAATGTGATCTGCAAACTTCTTTAATGATGAAGAGATGAAA 1012  
 QY 301 AsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSerSerProGlyPro 320  
 Db 1013 AATGTAAATGATTTCTTCAAAATGACCTTTCTGAATACCTGATCTATCTTCCCTGGACCA 1072  
 QY 321 ValGluIleLeuThrThrValThrGluSerValThrGlyIleSerArgAsnValSerThr 340  
 Db 1073 GTACAAATTTTAAACACAGTACAGAAATCAGTTACAGGTATTTCAAAATAATGTAAGCACT 1132  
 QY 341 AsnValPhePheIysGlnHisAspTyrIleIleGluPhePheAspTyrThrThrValLeu 360  
 Db 1133 AATGTGTCTTCAAGCAACATGATTAATCACTTAAGTTTTTGTATTAACATCACTGCTTGG 1192  
 QY 361 LysProSerLeuAsnPheThrAlaThrValLysValIleThrArgAlaAspGlyAsnGlnLeu 380  
 Db 1193 AAGCATCTCTCAACTTCAACGCACTGTGAAAGTAACTGCTGCTGATGGCAACCAACTG 1252  
 QY 381 ThrLeuGlnIleArgValAsnValValIleThrValIleThrGlnArgAsnTyrThrGlu 400  
 Db 1253 ACTTTTAAAGAAAGAAATTAATGTAGTCAATACAGTGACACAGAAACTATATACGAG 1312  
 QY 401 TyrThrSerGlySerAsnSerGlyAsnGlnLysMetGluAlaValIleGlnLysIleAsnTyr 420  
 Db 1313 TACTGAGCGGATCTTAACAGTGAATTCAGAAATGGAAGCTGTTCGAAATAATTAATAT 1372  
 QY 421 ThrValProGlnSerGlyThrPheIysIleGluPheProIleLeuGluAspSerSerGlu 440

Db 1373 ACTGTCCCAAGTGAACCTTTTAAGATTGAATTCCTCAATCTCGAGATTCACGTAG 1432  
 QY 441 LeuGlnLeuLysAlaTyrPheLeuGlySerIysSerSerMetAlaValHisSerLeuPhe 460  
 Db 1433 CTACAGTGAAGGCTTAATTTCTTGGTAGTAAAGACATGGCACTTCACTATGCTGTT 1492  
 QY 461 LysSerProSerIysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLysValIle 480  
 Db 1493 AAGTCTCTTAAGATGATCATATCACTCACTTAAACAGAGATGAAATATTAAGTGGGA 1552  
 QY 481 SerProPheGluLeuValValIleSerGlyAsnLysArgLeuLysGluLeuSerTyrMetVal 500  
 Db 1553 TCCCTTTTGAATGGTGGTGTAGTGGCAACAAACGATTTGAAGAGCTTATATGTATA 1612  
 QY 501 ValSerArgIleGlnLeuValAlaValIleGlyLysGlnAsnSerThrMetPheSerLeuThr 520  
 Db 1613 GATCCAGGGGACAGTGGTGGCTGTAGGAAACAAATTCAAACAAATGTTCTTTAACA 1672  
 QY 521 ProGluAsnSerTyrThrProLysAlaCysValIleValTyrTyrIleGluAspAspGly 540  
 Db 1673 CAGAAATATTTCTGACCTCCAAAGCTGTGTATTTGTATTAATTTGAAGATGATGGG 1732  
 QY 541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560  
 Db 1733 GAAATTAATAAGTATGTTCTTAAAAATTCCTGTTCACTGTTTAAAAATAAGATAAG 1792  
 QY 561 LeuTyrThrSerIysValIysAlaGluProSerGluLysValSerLeuArgIleSerVal 580  
 Db 1793 CTATATTTAGTAAAGTAAAGTGAACCTGAACCTGGAAGAGTCTCTTGTAGATCTCTGTG 1852  
 QY 581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600  
 Db 1853 ACACAGCTGACCTCCATAGTGGGATTTAGCTGTTGACAAAGTGTGAATCTGATGAAT 1912  
 QY 601 AlSerAsnAspIleThrMetGluAsnValValHisGluLeuGluLeuTyrAsnThrGly 620  
 Db 1913 GCCTTAATGATTAATCAATGAAATGTGTCATGAGTGGAACTTTAATACACAGCA 1972  
 QY 621 TyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTyrVal 640  
 Db 1973 TATTAATTTAGGCAATGATTAATGAAATTTCTTTGACGCTTTCAAGAAATGTGACTCTGGTA 2032  
 QY 641 LeuThrAspAlaAsnLeuThrIysAspTyrIleAspGlyValIleTyrAspAsnAlaGluTyr 660  
 Db 2033 TTACAGATGCAACCTCAAGAAAGATTAATTTAGTGTGATTAATGATGATGATGAT 2092  
 QY 661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680  
 Db 2093 GCTGAGAGTATTAAGAGAAATGAAGACATATTTGTAGATTTATCATGACTTTTCTTGG 2152  
 QY 681 GlySerSerProHisValArgLysHisPheProGluThrTyrIleThrLeuAspThrAsn 700  
 Db 2153 GGTGAAGTCCACATATCCGAAAGCAATTTCCAGACATGTGATTTGGCTGACACCAAC 2212  
 QY 701 MetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIleThrSerTyr 720  
 Db 2213 ATGGGTTACAGATTTTACAGAAATTTGAAGTACGTACCTGATTTCACTTCTTGG 2272  
 QY 721 ValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrThrProVal 740  
 Db 2273 GTGGCTACAGGTTTGTGATCTGAGGACCTGGGCTTGGACTTAACAATCACTCCAGTG 2332  
 QY 741 GluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArg 760  
 Db 2333 GAGCTCAAGCCCTTCAACCATTTTTCATTTTAACTTCCCACTCTGTATTCAGA 2392  
 QY 761 GlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrGluVal 780  
 Db 2393 GGTGAAGATTTGCTTGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2452  
 QY 781 LysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsn 800



Db 2453 AAGGTATCATTCAGAAAGTGAACAATTGATATTTCTAATGACTTCAAAATGAATAAAT 2512  
 QY 801 AAThGlyYHISgInGlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPhe 820  
 Db 2513 GCCACAGGCCACAGCAGACCTTCTGGTTCCAGAGAGAGATGGGCAACTGTTCTTTT 2572  
 QY 821 ProIleArgProThrHisIleuGlyGluIleProIleThrValThrAlaLeuSerProThr 840  
 Db 2573 CCATACAGGCCAACACATCTGGAGAAATTTCTATACAGTCAACACTCTTTCACCCACT 2632  
 QY 841 AlaSerAspAlaValThrGlnMetIleLeuValIleGluGlyIleGluIleuSerIle 860  
 Db 2633 GCTTCGATGCTCTACCCAGATGATTTTGTAAAGGCTGAAGAAATGAATAATCATAT 2692  
 QY 861 SerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuSerThrLeu 880  
 Db 2693 TCACAAATTCATCTTATTAAGCTTGACGACAAATAGGCTACAGGTAACCTGAAATCTTG 2752  
 QY 881 SerPheSerPheProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIle 900  
 Db 2753 AGTTTCTCATTTCTCTCTATACAGTCACTGCGAGAAAGATTCAAGATCACTGCAATT 2812  
 QY 901 GlyAspValIleuGlyProSerIleAsnGlyLeuAlaSerIleuIleArgMetProTyrGly 920  
 Db 2813 GAGGATGTTCTTGCTCTTCATCAATAGGCTTAGCCCTCATGTGATGGATGCTTATGGC 2872  
 QY 921 CysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThr 940  
 Db 2873 TGTGGTGAACAGAACATGATTAATTTTGTCCCAAAATTTTACATTTTGATTTCTGACT 2932  
 QY 941 LysIleLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPheMetArgGlnGly 960  
 Db 2933 AAAAAAACAACCTGACAGATTAATTGAAGAAAGACTCTTCAATTATGAGCAAGGT 2992  
 QY 961 TyrGlnArgGlnLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyr 980  
 Db 2993 TACCAAGAGAACTTCTCTATCAGAGGAGAGATGGCTTTTCAGTCTTTTGGGAATTTAT 3052  
 QY 981 AspProSerGlySerThrTyrLeuSerAlaPheValLeuArgCysPheLeuGluAlaAsp 1000  
 Db 3053 GACCTTCTGGAGGACTTGTTGTCTGACGTTTGTGTTTAAAGATGTTTCTTGAAGCCGAT 3112  
 QY 1001 ProTyrIleAspIleAspGlnAsnValIleuHisArgThrTyrThrTyrLeuLysGlyHis 1020  
 Db 3113 CCTTACATAGATTTGATCAGATGATGTTTACACAGAACATACACTTGGCTTAAAGACAT 3172  
 QY 1021 GlnLysSerAsnGlyGluPheTyrPheProGlyValArgValIleHisSerGluLeuGlnGly 1040  
 Db 3173 CAGAAATCCAAACGGTAAATTTTGGATCCAGAAAGATGATTCATGTGAGCTTCAAGGT 3232  
 QY 1041 GlyAsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArg 1060  
 Db 3233 GGGCAATAAAGTCCAGTAACATTAAGGCTTATATGTAATCTCTCCCGGGAATATGA 3292  
 QY 1061 LysTyrGlnProAsnIleAspValGlnGluSerIleHisPheLeuGlnSerGluPheSer 1080  
 Db 3293 AAGTACAGCCTTAACATTTGATGCAAGAGTCAATCATTTTGGAGTCTGAATTCAGT 3352  
 QY 1081 ArgGlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGly 1100  
 Db 3353 AGAGGAATTTCAACAATATTAATCTTAGCCCTTATTAATTCATATTCATCAGAGGGG 3412  
 QY 1101 SerProLysAlaLysGluAlaLeuAsnMetLeuThrTyrArgAlaGlnGlnGluGly 1120  
 Db 3413 AGTCTTAAGCCGAAGAGCTTGAATATGCTGACTTGGAGACCAAGAAAGAGTGGC 3472  
 QY 1121 MetGlnPheTyrValSerSerGluSerIleLysLeuSerAspSerTyrGlnProArgSerIle 1140  
 Db 3473 ATGCATTTCTGGGTGTCATCAGAGTCCAAACTTTCGACCTCCGCGACGCAACCTCCCTG 3532  
 QY 1141 AspIleGluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGlu 1160  
 Db 3533 GATATTGAAGTTGCAACCTATGCACTGCTCTCACACTTCTTCAATTTGAGACTTCTGAG 3592

QY 1161 GlyIleProIleMetArgTyrLeuSerArgGlnArgAsnSerLeuGlyGlyPheAlaSer 1180  
 Db 3593 GGATCCCAATATATGAGGTGGCTTAAGCAGCAAAAGAAATAGCTGGGTGTTTGCATCT 3652  
 QY 1181 ThrGlnAspThrThrValAlaLeuLysValAlaLeuSerGluPheAlaAlaLeuMetAsnThr 1200  
 Db 3653 ACTGAGATACCACTGTGGCTTTAAAGGCTGTGTGTAATTTGCAAGCCCTTAATGATACA 3712  
 QY 1201 GluArgThrAsnIleGlnValThrValThrGlyProSerSerProSerPro----- 1217  
 Db 3713 GAAAGCAAAATATCCAAATGACGTTGAGGAGGCTTAGCTTCAACAGTCTCTGTAAGTTT 3772  
 QY 1218 -----LeuAlaValGlnPro 1223  
 Db 3773 CTGATTGACACACACACCGCTTACTCTTACAGACAGACAGCTTCTGTGTACAGCCA 3832  
 QY 1224 MetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValVal 1243  
 Db 3833 ATGCGAGTTAAATATTTCCCAATATGTTTGGATTTGCTAATTTGTCAGCTCAATGTTGTA 3892  
 QY 1244 TyrAsnValIleAlaSerGlySerSerArgArgArgSerIleGlnAsnGlnGluAla 1263  
 Db 3893 TATTAATGTAAGGCTTCTGGGCTTCTTGAAGACCAAGATTCATCCAAATCAAGAGCC 3952  
 QY 1264 PheAspLeuAspValAlaValIleGluLeuAsnLysAspAspLeuAsnHisValAspLeuAsn 1283  
 Db 3953 TTTGATTTGATGCTTGCTGTAAAGAAATTAAGATGATCTCAATCATGTGATTTGAAT 4012  
 QY 1284 ValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluValAsnLeu 1303  
 Db 4013 GTGTGTAACAGCTTTTCCGGCCCGGATGAGATGCGATGCTCTTATGAAATTAACCTTA 4072  
 QY 1304 LeuSerGlyPheMetValProSerGluAlaIleSerLeuSerGluThrValIleLysVal 1323  
 Db 4073 TTAAGTGCTTATATGCTGCTTCAAGAACAAATTTCTGAGCGAGACAGTGAAGAAAGTG 4132  
 QY 1324 GluTyrAspHisGlyLysLeuAsnLeuTyrLeuAspSerValAsnGlnThrGlnPheCys 1343  
 Db 4133 GAATATGATCAGTGAAGAACTCAACTCTATTAGATTCTGTAAATGAACCCAGTTTGT 4192  
 QY 1344 ValAsnIleProAlaValArgAsnPhelyValSerAsnThrGlnAspAlaSerValSer 1363  
 Db 4193 GTTAATATCTCGTGTGGAACCTTAAAGTTTCAAAATACCAAGATCTTCAGGTGC 4252  
 QY 1364 IleValAspTyrTyrGluProArgArgGlnAlaValArgSerTyrAsnSerGluValLys 1383  
 Db 4253 ATAGGGAATTAATAGAGCCAAAGAGACAGCGGTGAGAAAGTTTACAACTCTGAAGTGAAG 4312  
 QY 1384 LeuSerSerCysAspLeuLysSerAspValGlnGlyCysArgProCysGluAspGlyAla 1403  
 Db 4313 CTGTCTCTCTGACCTTTTGCATGATGTCAGAGGCTGCCGTCCTTGTAGAGATGAGCT 4372  
 QY 1404 SerGlySerHisHisIleSerSerValIlePheIlePheCysPheLysLeuLeuTyrPhe 1423  
 Db 4373 TCAGGCTCCATCACAACACTTCAGCATTTTATTTCTGTTTCAAGTCTGTACTTT 4432  
 QY 1424 MetGluLeuTyrPhe 1428  
 Db 4433 ATGGAACCTTGGCTG 4447  
 RESULT 5  
 AX537612 4335 bp DNA linear PART 23-NOV-2002  
 LOCUS AX537612  
 DEFINITION Sequence 3 from Patent WO02070738.  
 ACCESSION AX537612  
 VERSION AX537612.1 GI:25269535  
 KEYWORDS  
 SOURCE  
 ORGANISM Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE 1

AUTHORS Schuh, A. and Ouwehand, W.  
TITLE Diagnosis and treatment of blood disorders  
JOURNAL Patient: WO 02070738-A 3 12-SEP-2002;  
Schuh, Andre (CA)

FEATURES  
SOURCE Location/Qualifiers  
1. .4335

/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
1. >4335  
/note="unamed protein product"

CDS

/codon\_start=1  
/protein\_id="CAD57497.1"  
/db\_xref="GI:25269536"  
/translation="MOGEPPLTAHLACVCTALAVAGPRFLVTAGIRPGAVNTI  
GVELLEHCPQVTKALKELTKASNLTVSVLEAGVPEKSGSKTLTSLPSNSADELY  
ELATYGRQDEILFNSNTRLSPEKRLISVFIQTDKALKRQEKREIVTLPSDFKRY  
KTSNLILIKDPSNLIQOWLSQSDLGISIKTQLSHPILIGDSIQVQVNDQTYOS  
FOVEBYVLPRKEVTLQTPLYCSNNSKHLNGITIAKYGKVKQDVLTLPSPFMWK  
KKNTTKFKINGSANFSFNDEMKVMSNGLSLYLDSPPGVEILITVTSVGTI  
SRNVTNVEFKQHDYITIEFDYTVLKPSLNTATVKTATRADQQLTLEBRNNVIT  
VTORNTVEYSGNSGNGOKMBAVOKIKVTVPOSGTPKIEPILEDSEELQKAYFGS  
KSNVAHSLFKSPSKTYIQLKTRDENIKVSPFELVYSGNKLKELSYMVVSGQLVA  
VGKQNSTMPSLTLPNSWTPKACVIVYIEDGELISVLKIPVOLVERKMKIUYMSKY  
KABPSEKVSLEISVTQDPSIVGIVAVKSVNLNANSNDITMENVEHLELJNTGYLGS  
MFNNSFAVQCEGMLVLDANLTKOYIDGYDNALVAREMEENHGIIDIDHSLGS  
SPHYRKHPEETWLDITNMGSRIVQEPFVTPDSITSWATGPFVISEDGLGTTTTP  
ELQAPQPEFTFLNPYSVIRGEFALETTFNYLKADTEYVYIIKSKDKDILMTSBE  
INATGHOOTLIVPESDGAIVLPIRPHLGIPIVTTALSPSTADAVTQMLVAEET  
EKYSQSLDLIDLNRLOSTLKITLSFSPNPVTGSEBVOITADIVLPGINGLAL  
IRMPYGGCEQNMNIFAPNIYILDYLTKKQLTDNLKSKALSFMGQVORBLVQREBG  
SFAFGNDPDSGTWSAFVLCLEADPYIDIDONLHRYTLLKQHSNGSFPMP  
GLRVHSELQGGNKSPTLLAYITSLIGRYKQPNIDVOESIHFLSESLQKAYFGS  
LALITLSSVSGSPKAKALNMLTMRABOEGMOPWVSSEKLSDSMPSPSIDENVA  
YALISHLPQFTSGEIPIMRWLSRQNSLGGFASDTQTLVAKLSSEFALMTERIN  
IQVTVTGSPSPVKFLIDTNRLILQTAELVAVQMANISANGFEPALCOLNVVN  
VKAGSSRRRSRSIONEAFDLDVAVKEDKLDNHDVNCISFGPGRSGALMELVN  
LSGPMVSEALSETVYKVEYDGHKLLYDSVNETOFVNIIPAVNFKVSNODAS  
VSIVDYBPRRQAVRSYNSVKKLSSCOLCSVQGRCPEDGASGSHHSSVIFPCKK  
LTYPMELML"

BASE COUNT 1295 a 838 c 922 g 1280 t  
ORIGIN

## Alignment Scores:

Pred. No.: 0 Length: 4335  
Score: 7317.50 Matches: 1426  
Percent Similarity: 98.75% Conservactive: 1  
Best Local Similarity: 98.69% Mismatches: 17  
Query Match: 99.58% Indels: 1  
DB: 6 Gaps: 1

US-10-020-095-4 (1-1428) x AK537612 (1-4335)

QY 1 MetGInglYProProLeuLeuThraAlaHisLeuLeuCyValCyThraAlaAlaLeu 20  
Db 1 ATGAGAGGAGCCACCGCTCTGACCGCGCCACCTCTCTCGTGTGACCGCGCGCTG 60  
QY 21 AlaValAlaProGlyProArgPheLeuValThraAlaProGlyIleIleArgProGly 40  
Db 61 GCCGTGGCTCCCGGCGCTCGGTTCTGTGTGACAGCCCGAGGATCATCAGCCCGAGGGA 120  
QY 41 AsnValThrIleGlyValGluLeuLeuGluHisCysProSerGlnValAlaThraValAla 60  
Db 121 AATGTGACTATTGGGAGTCTGTGACCTGACCTGCTTCTGACAGGTGACTGTAGAGCG 180  
QY 61 GluLeuLeuYsrThraAlaSerAsnLeuThraValSerValLeuGluAlaGluGlyValPhe 80  
Db 181 GAGCTGCTCAAGACAGCATCAACCTCATCTCTCTGCTGGAAGACAGAGAGTCTTT 240  
QY 81 GluYsglySerPheLeuYsrThraLeuThraProSerLeuProLeuAsnSerAlaAspGlu 100  
Db 241 GAAAGAGGCTCTTTAAGACACTTACTCTTCATCATCACTTGAAACAGTGCAGATGAG 300

QY 101 IleTyrGluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThr 120  
Db 301 ATTATGAGCTACGTGTAAACCGGACGATCCAGAGATGAGATTTATCTCTAATAGTACC 360  
QY 121 ArgLeuSerPheGluThrIleYsrGlyIleSerValPheIleGlnThrAspIleValLeuTyr 140  
Db 361 CGCTTACATTTAGACCAAGAAATATGTCGTTCAATCAACAGCAAGGCGCTTATAC 420  
QY 141 LysProLysGlnGluValLysPheArgIleValAlaThrLeuPheSerAspPheLysProTyr 160  
Db 421 AAGCCAAAGCAAGAGAGAGTTCCGATGTTACACTCTTCCAGATTTTAAAGCCTTAC 480  
QY 161 LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnIleTyrLeu 180  
Db 481 AAAACCTCTTAAACATTTCTTAAGACCCCAATCAATTTGATCCAAAGCTGGTTG 540  
QY 181 SerGlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIle 200  
Db 541 TCACAAACAAAGTATCTTGAGTCATTTCCAAACCTTTCAAGTATCTTCCATCCAAATA 600  
QY 201 LeuGlyAspTrpSerIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGln 220  
Db 601 CTGGTGACTGGTCTATTCAAGTTCAGATGAAATGACCAAGACATATTCATCATTTTACG 660  
QY 221 ValSerGluTyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSer 240  
Db 661 GTTTCGAATATGATATACCAAAATTTGAAGTACCTTGACAGACACATATATTTGTTCT 720  
QY 241 MetAsnSerLysHisLeuAsnGlyThrIleThraAlaLysTyrThrTyrGlyLysProVal 260  
Db 721 ATGAATTTCAAGCATTTAAATGATGATACCATCAGCGCAAGATACATATGGAAGCCAGTG 780  
QY 261 LysGlyAspValThrLeuThrPheLeuProLeuSerPheTrpGlyLysLysAsnIle 280  
Db 781 AAGAGACCTTACGCTTATATTTTACCTTATCTTTGGGAAAGAAAGAAATATT 840  
QY 281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluIleTyr 300  
Db 841 ACAAAAACATTTAAGATTAATGATGATCGCAAACTTCTCTTTATGATGAAAGATGAAA 900  
QY 301 AsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSerSerProGlyPro 320  
Db 901 AATGTATGATTTCTTCAATATGATGATCTTCAATACCTGATATCTTCCCTGAGCCA 960  
QY 321 ValGluIleLeuThrThrValThrGlnSerValThrGlyIleSerArgAsnValSerThr 340  
Db 961 GTAGAAATTTTAAACACAGTACAGATCATGTTACAGTATTTTCAAGAAATGATAGCACT 1020  
QY 341 AsnValPhePheLysGlnHisAspTyrIleIleGluPhePheAspTyrThrThraValLeu 360  
Db 1021 AATGTGTTCTTCAAGCAACATGATTTACATCATGATTTTTCATTATACTACTGCTTG 1080  
QY 361 LysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsnGlnLeu 380  
Db 1081 AAGCATCTCTTCAATTCACAGCAGCTGTGAGATGATCTGTCTGATGCAACCACTG 1140  
QY 381 ThrLeuGlnGluArgArgAsnAsnValIleIleThrValThrGlnArgAsnTyrThrGlu 400  
Db 1141 ACTCTTGAAAGAAAGAAATATGATGATATACAGTGCACAGAGAACTATACCTGAG 1200  
QY 401 TyrTrpSerGlySerAsnSerGlyAsnGlnLysMetGlnAlaValGlnLysIleAsnTyr 420  
Db 1201 TACTGAGCGGATCTTAACGTGGAATTCAGAAATGGAAGCTGTTTCAGAAATTAATAT 1260  
QY 421 ThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGluAspSerSerGlu 440  
Db 1261 ACTGCCCCCAAGTGGAACTTTTAAAGATTTCCCATCTCGGAGGATTCACAGTACG 1320  
QY 441 LeuGlnLeuValAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSerLeuPhe 460  
Db 1321 CTACAGTTGAAGGCTATTTCTTGTGATGAAAGTGCAGTGCAGTTCATCTGTTT 1380  
QY 461 LysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLysValGly 480

D	1381	AAAGTCTCTAGTAAGACATACATCCAACTTAAAAACAAGAGATGAAAATATTAAGGTGGGA	1440
Q	481	SerPropegluLeuValValSerGlyAsnGlySerGluLeuSerTyMetVal	500
D	1441	TGCGCTTTGAGTTGGTGGTTAGTGGCAACAAACGATGAAGAGATTAAAGTATATGGTA	1500
Q	501	ValSerArgGlyGlnLeuValAlaValGlyGlnAsnSerThMetPheSerLeuthr	520
D	1501	GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAAAATAATTCACACATGTCCTTTTAA	1560
Q	521	ProGluAsnSerTyrThrProValAlaCysValIleValTyrTyrIleGluAspArgI	540
D	1561	CCAGAAATTTCTTGACATCCAAAGACCTGTGTAAATGGATTAATTAATGAAGATATGG	1620
Q	541	GluIleIleSerAspValLeuGlyIleProValGlnLeuValPheLeuAsnGlyIleG	560
D	1621	GAAATTATTAAGATGTTCTTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATTAAGTTAAG	1680
Q	561	LeuTyrTyrPheSerValIleGluProSerGluValSerLeuArgIleSerVal	580
D	1681	CTATATTTGGAGTAAGTAAGTAAAGCTGAACCATCTGAAAGTCTCTCTTAAAGATCTCTGG	1740
Q	581	ThrglnProAspSerIleValGlyIleValAlaValAlaPlySerValAsnLeuMetAsn	600
D	1741	ACACAGCTTGACTCCATAGTTGGAGTGTGAGCTGTGTCACAAAGTGAATCTGATGAT	1800
Q	601	AlaSerAsnAspIleThrMetGlyAsnValAlaIleSGluLeuGluLeuTyrAsnThrGly	620
D	1801	GCCTTATATGATATTAACAATGGAATAATGTGGCCATGATGATCGAATCTTATTAACACAGA	1860
Q	621	TyrTyrIleGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTyrVal	640
D	1861	TATATTTTAGGATGTCATGATTAATCTTTTGACAGCTTTCAGAGATGTGACTCGGGTA	1920
Q	641	LeuThrAspAlaAsnLeuThrIleAspTyrIleAspGlyValTyrAspAsnAlaGluTyr	660
D	1921	TTGACAGATGCACAAACCTCAGCAGAGATTATATGTATGTGTTTAAACAAGCAGAAAT	1980
Q	661	AlaGluAspPheMetGlyGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu	680
D	1981	GCTGAGAGCTTTATGAGAGAAAATTAAGACATATGTAGATATTCATGACCTTTCTTTG	2040
Q	681	GlySerSerProHisValArgIleHisAspProGlnThrTyrIleTyrLeuAspThrAsn	700
D	2041	GGTAGCAGTCCACATGTCGGAAGCATTTTCCAGAGACTTGGAATTTGGCTGACACCAAC	2100
Q	701	MetGlyTyrPheGlyIleTyrGlnGluPheGluValThrValProAspSerIleThrSerTyr	720
D	2101	ATGGGTTCCAGGATTTACCAAGAAATTTAAGTAACCTGATTCATACATCTTG	2160
Q	721	ValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrProVal	740
D	2161	GTGGCTACTGGTTTGTATCTCTAGGACCTGGGTCTTGGACATAACAACATCTCCAGTG	2220
Q	741	GluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArg	760
D	2221	GAGCTCCAGGCTTCCAAACATTTTCAATTTTTTTGATCTTCCCTACTCTGTATACGA	2280
Q	761	GlyGlnGluPheAlaLeuGluIleThrIlePheAsnTyrLeuAspAlaThrGluVal	780
D	2281	GGTGAAGAATTTGCTTGGAAATACTATATTCATTTTGAAGAAGGCCACTGAGGTT	2340
Q	781	LysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsn	800
D	2341	AAAGTAATCATTTGAAGAAAAGTACAAATTTGATATCTTAAGACTTCAAAAGAAATAAT	2400
Q	801	AlaThrGlyHisGlnGlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPhe	820
D	2401	GGCAAGGCGACACACAGACCTTCTGGTCCCACTGAGATGGGCAACTGTTCTTTT	2460
Q	821	ProIleArgProThrHisLeuGlyGluIleProIleThrValThrAlaLeuSerProThr	840

Db	2461	CCCAATCAGGCCAACACACACTCTGGAGAAATTCCTTATCACAAGTCACAGCTCTTACCCACT	2520
QY	841	AlaseaSpAlaValThrGlnMetIleuValIysalagIlyIleGluYsSerYr	860
Db	2521	GCTTCGTAGCTGTACCCACAGATATTTTAAAGGTGAAGAAATGAAAATTCATAT	2580
QY	861	SerGlnSerIleIleuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuYsThrLeu	880
Db	2581	TCACAAATCCAACTTAATAGACTTCACTGACAAATAGGCTACAGAGATACCTGAAAACTTTG	2640
QY	881	SerPheSerPheProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIle	900
Db	2641	AGTTTCTCATTTCTCTCTTAATACAGTACTGGCAGTGAAGAAGTTCAATCACTGCATTT	2700
QY	901	GlyAspValIleuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGly	920
Db	2701	GGAAGATGTTCTGGGCTTCATCAATAGGCTTACCTCATTTGATTCGGATGCTTATGGC	2760
QY	921	CysGlyGlyGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThr	940
Db	2761	TGTGTGTGAACGAACAATGATTAATTTTGCTCCAAATATTTCACATTTTGGATTACTGACT	2820
QY	941	LysIlyslYsGlnLeuThrAspAsnLeuIlyslYsAlaLeuSerPheMetArgGlnGly	960
Db	2821	AAAAAGAAACAACCTGACAGATTAATTTGAAAAAAAAGCTCTTCAATTTATAGAGCAAGGT	2880
QY	961	TyrGlnArgGlnLeuLeuTyrGlnArgGlnAspGlySerPheSerAlaPheGlyAsnTyr	980
Db	2881	TACCAAGAGAACTTCTCTATCAGAAGGAAATGGCTCTTCAAGTCTTTGGGAATTAT	2940
QY	981	AspProSerGlySerThrTyrPLeuSerAlaPheValLeuArgCysPheLeuGluAlaAsp	1000
Db	2941	GACCTTCTGGAGACACTTGGTGTCTCAGCTTTTGTTTAAAGATGTTTCTTGAAGCCGAT	3000
QY	1001	ProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTyrPLeuYsGlyHis	1020
Db	3001	CTTACATAGAGATATTGATCAGAATGTGTTCACAGAACATACACTTGGCTTAAAGGACAT	3060
QY	1021	GlnYsSerAsnGlyGluPheTyrAspProGlyArgValIleHisSerGluLeuGlnGly	1040
Db	3061	CAGAAATCCACGGTGAATTTTGGGATCCAGGAAGAGTATTCATAGACCTTCAGAGT	3120
QY	1041	GlyAsnYsSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArg	1060
Db	3121	GGCAATAAAGATCCAGTACACTTACAGCCTATATTGTAATCTTCTCTCGGATATATGA	3180
QY	1061	LysTyrGlnProAsnIleAspValGlnGlnSerIleHisPheLeuGlnSerGluPheSer	1080
Db	3181	AAGATATCGCTTAACATTTGATGTCCAGAGCTATTCATTTTTTGGAGTCTGAATTCCAGT	3240
QY	1081	ArgGlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGly	1100
Db	3241	AGAGGAATTTAGACAAATTAATACCTAGCCTTTAACTTAATAGCATGTCTACAGTGGGG	3300
QY	1101	SerProYsAlaYsGluAlaLeuAsnMetLeuThrTyrParGlnIleGlnGlnGlyGly	1120
Db	3301	AGTCTTAAGGGAAGGAAGCTTTGAATATGTCTGACTTGGAGACGAGAACAGAAAGGTGGC	3360
QY	1121	MetGlnPheTyrValSerSerGluSerIlyLeuSerAspSerTyrGlnProArgSerIleu	1140
Db	3361	ATGCAAATTTCTGGGTGTATCAGAGTCCAAACTTTCTGACTCTCGGACGCCAGCTCCCTGT	3420
QY	1141	AspIleGlyValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGlu	1160
Db	3421	GATATGTGAAGTTGACGCTATGCACTGTCTCAACAATCTTTCACATTTCAACACTTCTGAG	3480
QY	1161	GlyIleProIleMetArgTyrPLeuSerArgGlnArgAsnSerLeuGlyGlyPheAlaSer	1180
Db	3481	GGATTCCTCAATTAATAGAGGTGCTACAGCAGGAAAGAAATAGCTTGGGGTGTTTGCACT	3540
QY	1181	ThrGlnAspThrThrValAlaLeuIlyslYsAlaLeuSerGluPheAlaAlaLeuMetAsnThr	1200
Db	3541	ACTGAGGATACACTGTGGCTTTAAAGGCTGTGTCTAAATTTGCAAGCCTTATAGAAATCA	3600

QY	1201	GIUAQTHAENIIIEGLIVAlThrValThrGlyProSerSerProSerPro-----	1217
Db	3601	GAAGAACAATATCAAGATGACGTGACGGGGCTTACTCAACAAGTCTGTAAAGTTT	3660
QY	1218	-----LeuAlaValAlaGlnPro	1223
Db	3661	CTGATTGACACACACACCGCTTACTCTTACAGACGAGACGCTTGCTGTGTACAGCA	3720
QY	1224	MetAlaValAsnIIISerAlaAsnGlyPheGlyPheAlaIIeCyGlnLeuAsnValVal	1243
Db	3721	ATGGCAGTTAAATATTCGCAAAATGGTTTGGATTGTGATTGTTCAGCTCAATGTTGTA	3780
QY	1244	TYAaenValIlysaISerGlySerSerArgArgArgSerIIeGlnAsnGlnAla	1263
Db	3781	TRTATATGGAAGGCTTCTGGGCTTCTTGAGACACAAATCTATCCAAATCAAGAACCC	3840
QY	1264	PheAspLeuAspValAlaValIlysgLysnIlysaAspLeuAsnHisValAspLeuAsn	1283
Db	3841	TTTGATTAGATGTTGCTGTAAATAAATAAAGATGATCTCAATCATGTGATTGAT	3900
QY	1284	ValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGlnValAsnLeu	1303
Db	3901	GTGTGTACAAGCTTTTCCGGCCCGGGTGGAGGTGCATGGCTCTTATGGAGTTAACCTA	3960
QY	1304	LeuSerGlyPheMetValProSerGlnAlaIIeSerLeuSerGluThrValIlysgVal	1323
Db	3961	TTTAATGGCTTTATGTGTGCTTACAGAACATTTCTCTGACGACACAGTAAAGAACTG	4020
QY	1324	GluTYAAPHISGLIYLYsLeuAsnLeuTYrLeuAspSerValAsnGluThrGlnPheCys	1343
Db	4021	GAATGTGATCATGGAATACTCAACTTATTAGATCTGTAAATGAACCACTTGTGT	4080
QY	1344	ValaenIIeProAlaValArgAspPheIyValSerAsnThrGlnAspAlaSerValSer	1363
Db	4081	GTTAAATATTCCTCGTGTGAAACTTTAAAGTTCAATTAACCAAGATGCTTCAGTGC	4140
QY	1364	IleValaAspTYrTYrGluProArgArgGlnAlaValArgSerTYAaenSerGluValIlysg	1383
Db	4141	ATAGTGAATTTACTATGACCCAAAGAGACAGCGCGTGAAGAATTCAACTCTGAAGTGAAG	4200
QY	1384	LeuSerSerCYaAspLeuCYaSerAspValGlnGlyCYaArgProCYaGluAspGlyAla	1403
Db	4201	CTGTCTCTCTGTGACCTTTGGCAGTATGTCAGGGGCTGCCCTCTTGGAGATGAGACT	4260
QY	1404	SerGlySerHisHisIISerSerValIIepheIIePheCysPheIyLeuTYrPhe	1423
Db	4261	TCAGGCTCCACATCACTCTTCAGTCAATTTTATTTCGTATTCAGACTTCTGTACTTT	4320
QY	1424	MetGluLeuTYrLeu	1428
Db	4321	ATGGAACTTTGGCTG	4335
RESULT 6			
LOCUS	AX534936	4761 bp	DNA linear PAT 22-NOV-2002
DEFINITION	Sequence 3 from Patent WO02070696.		
ACCESSION	AX534936		
VERSION	AX534936.1	GI:25261478	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens (human)		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE			
AUTHORS	Schuh, A. and Sutherland, R. D.		
TITLE	Cd109 nucleic acid molecules and polypeptides and methods of use		
JOURNAL	Patent: WO 02070696-A 3 12-SEP-2002;		
	Schuh, Andre (CA); Sutherland, Robert D. (CA)		
FEATURES	Location/Qualifiers		
source	1..4761		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		

CDS

[/db\\_xref="taxon:9606"](#)  
113 . 4450  
/note="unnamed protein product"  
/protein\_id="CAD57232.1"  
/db\_xref="gi:25261479"  
/translation="MGQPLLTAAHLCLCVCTAALAVAPGPELVTAIPGIIIRGGNVITL  
GVELLHCPQVTVKALIKLTKASNLTVSVLEAGVFKSPTLTLPSPINSADEIIV  
ELRATYGTQDEILFNSNTRLSPEYTRISVPTQTRALKYKPOEVKFRVYLPSPKPY  
KTSNLILIKDPSNLIQWMSQSDILEYKTFQSHPLIGDMSIQVQVDDTYQD  
FOVEYVLPKEEVLQTPLYCSNMSKLNGTITAKYTGKSGKVDVLTLPPLSPWG  
KKNITKFKINGSNANFSFNDENKNVMSNGLEIYDLSKPEVELITVTESVTG  
SRANSTVYFPKHDIYIEFPDYTVLPKSLNFTVATKADGNQLTEBRNNVVI  
VTQNTNTEWYSGNSGNQKMAVOKIVTVATGQSTFKPEPILDESDLOKAYELG  
KSNNAVHSLFKSPSKTYIQLKTRDENIKVGSPEELVYSGNKLKELSTMYVSRQVAV  
VGKNSTWFSLTLPENSMTPKACVIVYIEDGEIISVDLKIPLVQVFNKIKLIMSK  
KAEBSEKSLRISVTQPSIVGIVADVKSVMANASNDITMENNVHELNTGTYLID  
MFNNSPAVQECGLMVLTDALITKDYIDGVNDMEYERFMEENEGHIVDIDHSIGS  
SPHRKHPETMIWLDTNMGSRIVIOEFPVYVPSDITSVVAGPVSIEDLGITTPPE  
ELQAPQPPPIFLNIPSYVINGEERALEITIPNYLKADTEYVYIEKSKPDIIMSN  
INAGHQDILLVPESDGAIVLPFIRPHLEGIPIVVALSTPASDAVYOMLVKABGI  
EKASQSIILLDLTDNRLOSTLKTLSFSPPEVTVGSRVQITADGVLPISNGIASI  
IRMYGCGEOMNINAPNIYILDYLTAKKQJTDMLKELKASFMQOGRBELLYREDP  
SFSGFNVDPGSGTMAFVLCFLEADPYDIDONLHRTYVILKQKNGSGEFDVPI  
GRVHSELQGGKSPVLTAVIYVLSILGYRQVGNPIVQESIHFLSEFSRGIIDNYI  
LALITVALSSGSKPAKELAMLMLTRABQEGMFWMSSEKSDSNQPRSDILEVA  
YALISHLPQFQTSIGTPIETMWLSKORNSLGSFASDTQVYALKALSEPALMNTERT  
IQVNTGPSSPSPVKFLIDITHNRLILQTELVAVQPAVNIISANGFALIQLVNVYV  
VKAGSSRRRSRRIIONOAFDLIDVAKENKDLINVDLVNCTSPGSGKVALMEVNI  
LSGFWVSEPAISLSETVKAYEDRGKILYDSVNEQFCVNI PAVNRFKSVNTQDA  
VSIVDYVPEPQARVRSYNSVEKLSCLQSDVQGCRCPEDGASGSHHSVIFIFCFP  
LLYMEMLT"

BASE COUNT 1419 a 913 c 1016 g 1413 t

ORIGIN

Alignment Scores:

Pred. No.:	0	Length:	4761
Score:	7317.50	Matches:	1426
Percent Similarity:	98.75%	Conservative:	1
Best Local Similarity:	98.69%	Mismatches:	1
Query Match:	99.58%	Indels:	17
DB:	6	Gaps:	1

US-10-020-095-4 (1-1428) x AX534936 (1-4761)

QY	1	MetGngIyProProlLeuLeuThrAlaIahIsleuLeuCyValCysThrAlaAlaLeu	20
DB	113	ATGCAAGGGCCCAACGGCTCTGACCGCGGCCCAACCTCCTCGTGGTGACCGCGCGCTG	172
QY	21	AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGly	40
DB	173	GCCGGGCTCCCGGGGCTCGGCTTCTGGTGACACCCCGAGGATCATCAGCGCCGGAGGA	232
QY	41	AsnValThrIleGlyValIleGluLeuLeuGlnHisCysProSerGlnValThrValIleAla	60
DB	233	AATGTGACTATTGG3GAGACTTCTGGAAACATGCGCTTCAACAGTACGTGAAGCG	292
QY	61	GluLeuLeuValThrAlaSerAsnLeuThrValSerValleuGlnIleGluGlyValPhe	80
DB	293	GAGCTGCTCAAGACAGATCAACTCACTCTCTGTCTCGGAAGCAGAGAGACTCTT	352
QY	81	GluIlyGlySerPheIlyThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGlu	100
DB	353	GAAAGGCTCTTTTAAAGACACTTACTTTCATCACTCACTCGAACAGTACAGATAG	412
QY	101	IleIlyrGluLeuAsgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThr	120
DB	413	ATTATATGAGCTACCTGTAAACCGAGCGATCCAGATGAGATTATATCTTAATATGACC	472
QY	121	ArgLeuSerPheGluThrIlyArgIleSerValPheIleGlnThrAspIlyAlaLeuIlyr	140
DB	473	CGCTTATCATTTGAGACCCAGAGAAATTCGTCTTCACTTCAACAGACAGAGGCTTATAC	532

QY	141	LYSPROLYSGINGLUVALIYSPHEARTGLEVALTHIRLEPNSESRASPHELYSPOTYR	160
DB	533	AAGCCAAAGCAAGATGAGTGGCATTTGGCATTTTACACTCTTCTCGATTTTAAAGCCTTAC	592
QY	161	LYSTHRSERLEASNIIELEUUILEYSAPPOLYSSERASPLEUILEGLINDINTHLEU	180
DB	593	AAAACTCTTTAAACATCTCATTAAGGACCCCAATCAAAATTTATTCACAGAGGTGG	652
QY	181	SERGINGINSESRASPLEUGLYVALIISESERYTHRPHEGLILEUSERSERHISPROILE	200
DB	653	TCACAAACAAGATCTTGATGATCATTTCCAAAACITTTGACGTATCTTCCATTCATTA	712
QY	201	LEUGIYASPTIPSERIEGLINVALGIVALAANASPGINTHRTYRTRYGLINSEPHGIN	220
DB	713	CTTGGGACTGGCTCATTCAGATTCAAGATGACAGACATATTAATCAATCATTTGCG	772
QY	221	VALSERGIUTYRVALLEUPROLYSPHEGLUVALITHIRLEGINTHPROLEUTYRYSER	240
DB	773	GIITTCAGAAATATGATATTCACAAAATTTGAAGACTTTGCACACACATTAATATGTTCT	832
QY	241	MECAENSERLEYSHISLEUASNGIYTHRIETHPALALSYTYRTRYGLIYASPROVAL	260
DB	833	ATGAATTCATACCATTTAAATGTCATACACGCGAAGATATACATATGGAGACCACTG	892
QY	261	LYSGIYASPAVALTHIRLEUTHRPHLEUPROLEUSERPHETIPGLIYSELVYASNIIE	280
DB	893	AAAGGAGACGTAAACGCTTACATTTTATCCTTATCTTTGGGAAAGAAATAATATT	952
QY	281	THIRYSTRPHLYSRIIEASNGIYSERIALANPNHSESRPHASNAHSPGLUIMETYS	300
DB	953	ACAAAACATTTAAAGATTAATGATCTGCAAACTTCTTTAAAGATGAAGAAAGATGAAA	1012
QY	301	ASNVALMELASPSERSESRASNGIYLEUSERGIUTYRLEUASPLEUSERSESRPROGLI	320
DB	1013	AATGTAATGATTTCTTCAAAATGACATTTCTGAAATACCTGATCTATCTTCCCTGGACA	1072
QY	321	VALGIUILEUTHRTHRYVALTHIRGUSERVALTHIRGIYIIESESRASNAVALSERTH	340
DB	1073	GTAGAAATTTTAAACACAGTGACACAGATCAGTTACAGGATTTTCAAGAAATGATGAC	1132
QY	341	ASNVALPHEPHLYSGINHISASPTYRIIEGLEUPHIEHASPITYRTHRYVALLEU	360
DB	1133	AATGGTCTTTCACACACATGATTAACATATGATGATTTTGAATTAATACACTGCTCTGG	1192
QY	361	LYSPROSERLEUASNPHERTHRALATHRYVALIYSVVALTHIRARVALASPGIYASNGINLEU	380
DB	1193	AMGCATCTCTCAACTTCACACCCACTGTGAAGGTACTCGTGTGATGGCAACCACTG	1252
QY	381	THIRLEUGIUGIUAIRGARASNAANVALIIEHTRVALTHIRGINARASNTYRTHIRGLU	400
DB	1253	ACTCTTGAAGAAAGAAATTAATGATCATTAACAGTGACACAGAGAAACATATCTGAG	1312
QY	401	TYRTTPSERGIYSESRASNSERGIYASNGINLYSMETGLUVALAIGLHLYSRIIEASNTYR	420
DB	1313	TACTGAGAGGGATCTTACACAGTGGAAATCAGAAAATGGAAGCTGTTCAAGAAATATTAATAT	1372
QY	421	THIRVALPROGINSESGIYTHRPHELYSIIIEGIUPHIEPROILEUGIUNASPSERSEIU	440
DB	1373	ACTGTCCCCCAAGAGAACTTTTAAGATGATATTCACCAATCTTGAGAGATTTCCAGTAG	1432
QY	441	LEUGINLEUYVALIYRPHLEUGIYSERIYSESRSEMETALAVAIHISESERLEUPHE	460
DB	1433	CTACAGTTGGAAGGCTTATTTCTTGATGTAATAAGTATACATGGCATCATAGTCTGTTT	1492
QY	461	LYSSERPROSERLYSETHRYRIIEGLINLEUYSTRIRGASPGIUNANIIELYVALIGLY	480
DB	1493	AAGTCTCTTAAGTAACATATCATTCACATTAAMACAGAGATGAAAATATTAAGGTGGGA	1552
QY	481	SERPROPHAGIULEUVALIYASERGIYASNUYSATGLIYSGIUNLEUSERTYRMEETAL	500
DB	1553	TGCGCTTTTGAGTGTGTGTTAGTGGCAACAAACATTAAGAGAGTTAAGCATTAATGGTA	1612
QY	501	VALSERARGIYGLINLEUVALAVALIGIYSGIINSESRTHMETPHESEIRLEUTHR	520

D	1613	GTATCCAGGGACAGTGGTGGCTGGTGGAGAAAACAAATTCACAAATGTTCTTTTAA	1672
Q	521	ProGluAsnSerTrpThrProValAlaCysAlaIleValTyrTrpIleGluAspArg	540
D	1673	CCAGAAATTCCTGGACTCCAAAAGCCGTGTAAATGTGATTTATTTGAAGATGAATGG	1732
Q	541	GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys	560
D	1733	GAAATTAATAGTGAATGTTCTTAAAAATTCCTTCAGCTTGTTTAAAAATTAAGTAAAG	1792
Q	561	LeuTyrTrpSerLysValLysAlaGlnProSerGlnLysValSerLeuArgIleSerVal	580
D	1793	CTATATGGAGTAAAGTAAAGCTGAACCATCTGAGAAAGCTCTTAGATCTCTGTG	1852
Q	581	ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn	600
D	1853	ACACAGCTGACTCCATAGTGGATGTGTACTGTGTGACAAAGTGTGAATCTATGAT	1912
Q	601	AlaSerAsnAspIleThrMetGluAsnValValHisGluLeuGluLeuTyrAsnThrGly	620
D	1913	GCTCTTAATGATATTAATCATAGAAAATGTGTCCATGAGTTGAACTTTAATACAGGA	1972
Q	621	TyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnLysCysGlyLeuTrpVal	640
D	1973	TATTATTAGCAGATGTCATGAATTCCTTTTGCAGTCTTTCAGAAATGTGACTGTGGGTA	2032
Q	641	LeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyrAspAsnAlaGluTyr	660
D	2033	TTGACAGATGCAAACTTCACAGAAAGATTAATATGATGATGTTAAATGCAATGCAAAATAT	2092
Q	661	AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu	680
D	2093	GCTGAGAGGTTTATGAGAGAAAATGAAAGACATATTTGATGATATTCATGACTTTTCTTTG	2152
Q	681	GlySerSerProHisValArgLysHisPheProGluThrTrpIleTrpLeuAspThrAsn	700
D	2153	GGTAGCAGTCCACATGCTCCGAAAGCATTTTCCAGACCTTGGAATTTGGCTTAGACACCAAC	2212
Q	701	MetGlyTyrTrpArgIleTyrGlnGluPheGluValThrValProAspSerIleThrSerTrp	720
D	2213	ATGGATTCACAGATTTTACCAAGAAATTTGAAGTAACTGTACTGATTTATCACCCTTCTGG	2272
Q	721	ValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrTrpThrProVal	740
D	2273	GTGGCTACTGATTTGTGATCTCTGAGGACCTGGGCTCTGGACCTAACCACTACTCCAGTG	2332
Q	741	GluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArg	760
D	2333	GAGCTCCAGGCTTCCAAACCATTTTTCATTTTGTGATCTTCCCTACTCTGTTATCAGA	2392
Q	761	GlyGlnGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrGluVal	780
D	2393	GGTACAGAAATTTGCTTTGAAAATCACTAATTTCAATATTGAAAAGATGCCACAGAGTT	2452
Q	781	LysValIleIleGlyLysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsn	800
D	2453	AAGGAATCATTTGAAAAGTGAACAAATTTGATTTCTAATGACTTCAAAATGAAAATAAAT	2512
Q	801	AlaThrGlyHisGlnGlnThrLeuLeuValProSerGluLysGlyAlaThrValLeuPhe	820
D	2513	GCCACAGGCCACGACGAGACCCCTCTGTGTTCCAGTAGAGATGGGGCACGTTCTTTT	2572
Q	821	ProIleArgProThrHisLeuGlyGluIleProIleThrValThrAlaLeuSerProThr	840
D	2573	CCCATTCAGGCCACACACATCTGGGGGAAATTCCTATCAAGTCAACGCTCTTTCACCCACT	2632
Q	841	AlaSerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIleGlyLysSerTyr	860
D	2633	GCTTCTGATGCTGCACCCACGATGATTTTGTATAGGCTGAAGAAATGAAAAATCATAT	2692
Q	861	SerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeu	880

Db 2693 TCACAAATCCATCTTATTAGACTGACAAATAGGCTACAGAGTACCCTGAAAACTTTG 2752  
 Qy SerPheSerPheProAsnThrValThrGlySerGluArgValGlnIleThrAlaIle 900  
 Db 2753 AGTTTCATCTTCTCTCAATACAGTACCTGGCAGTAAAGATTCAGATACCTGCAATT 2812  
 Qy GlnAspValIleGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProGlyGly 920  
 Db 2813 GGAATGTTCTTGGCTCTTCATCAATGGCTTACCTCATTCATTCGATTCGGATTCATGGC 2872  
 Qy CysGlnGluGlnIleAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThr 940  
 Db 2873 TGTGTGAACGAAACATGATTAATTTGTCTCAATATTTATTCATTTGGATTATCTCACT 2932  
 Qy LysIleLysGlnLeuThrAspAsnLeuGlyLysValAlaLeuSerPheMetArgGlnGly 960  
 Db 2933 AAAAAGAAACAACCTGACAGATTAATTTGAAAGAAAAGCTCTTTCAATTTATAGGCAAGCT 2992  
 Qy TyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyr 980  
 Db 2993 TACCAGAGAGAACTTCTCTATACAGGGAGATGGCTCTTCAGTGTCTTTGGCAATTAT 3052  
 Qy AspProSerGlySerThrTyrLeuSerAlaPheValLeuArgCysPheLeuGluAlaAsp 1000  
 Db 3053 GACCCCTTCTGGAGACACTGGTGTGCAGCTTTGTTCATAGATGTTCCTTGAAGCCGAT 3112  
 Qy ProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTyrPheLeuGlyHis 1020  
 Db 3113 CTTTACATAGATTTGATCAGAAATGTGTTACAGAAACATACATTCCTGCTTAAAGCAAT 3172  
 Qy GlnLysSerAsnGlyGluPheTyrAspProGlyArgValIleHisSerGluLeuGlnGly 1040  
 Db 3173 CAGAAATCCAAACGGTGAATTTTGGGATCCAGAAAGATGATTCATAGAGCTTCAAGGT 3222  
 Qy GlnLysLysSerProValThrLeuThrAlaTyrIleValThrSerLeuGlnGlyTyrArg 1060  
 Db 3233 GGCATATAAAGCTCAGTAACACTTACAGCCATATTTGATCTTCTCCGGAATATAGA 3292  
 Qy LysTyrGlnProAsnIleAspValGlnGluSerIleHisPheLeuGluSerGluPheSer 1080  
 Db 3293 AAGTATCAGCTTACATTCATGATGTGCAGAGTCAATTCATTTTGGAGCTGAAATCAGT 3352  
 Qy ArgGlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGly 1100  
 Db 3353 AGAAGAAATTCAGCAATTAATACCTAGCCCTTAATTAATGATTCATGCTGAGGG 3412  
 Qy SerProLysAlaLysGluAlaLeuAsnMetLeuThrTyrPheAlaGlnGlnGlyGly 1120  
 Db 3413 AGTCTTAAAGCGAAGGAGCTTTGAATATGCTGACTTGGAGAGCAAGAAAGGTGGC 3472  
 Qy MetGlnPheTyrPheLysSerGluSerIleLeuSerAspSerTyrGlnProArgSerLeu 1140  
 Db 3473 ATGCAATTTCTGGGTGTATCATCAGAGTCCAAACTTTCGATCTCTGACGACCAAGCTCCCTG 3532  
 Qy AspiIleGluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGlu 1160  
 Db 3533 GATTTGAAGTTGACGCTTAAGCACTGCTCTCACAATCTTACACAATTCAGACTTCAGAG 3592  
 Qy GlnTyrIleProLysMetArgTyrPheLysSerArgGlnArgAsnSerLeuGlyIlePheAlaSer 1180  
 Db 3593 GGAATCCCAATTAATGAGGTGCTTAAGCAGCGCAAGAAATAGCTTGGGTGGTTTGGCATCT 3652  
 Qy ThrGlnAspThrThrValAlaLeuLeuGluAlaLeuSerGluPheAlaAlaLeuMetAsnThr 1200  
 Db 3653 ACTAGAGATACACTGTGGCTTTAAAGCTCTGTGTAATTTGACAGCCCTTAATAGAACATA 3712  
 Qy GlnArgThrAsnIleGlnValThrValThrGlyProSerSerProSerPro----- 1217  
 Db 3713 GAAAGCAAAATATCAAGTACCGTGAAGGGGCTTACCTACCAAGTCCCTGTAAGTTT 3772  
 Qy 1218 -----IleuAlaValValGlnPro 1223  
 Db 3773 CTGATGACACACAAACCGCTTACTCTCTTACAGACAGAGCTTGTGCTGTGTACAGCCA 3832

Qy 1224 MetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValIle 1243  
 Db 3833 ATGCGATTATATATTTCCGCAATGGTTTGGATTTGCTATTGTCAGGCTCAATGTTGTA 3892  
 Qy 1244 TyrAsnValIleAlaSerGlySerArgArgArgSerIleGlnAsnGlnGluAla 1263  
 Db 3893 TATATGTGAAGGCTTCTGGGTCTTCTAGAAAGCGAAGATCTTATCCAAATCAAGAGCC 3952  
 Qy 1264 PheAspLeuAspValAlaValLysGluAsnLysAspAspLeuAsnHisValAspLeuAsn 1283  
 Db 3953 TTTGATTTACATTTGCTGTAAAGAAATTAAGATGATCTCAATCATGTGGATTGGAAT 4012  
 Qy 1284 ValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluValAsnLeu 1303  
 Db 4013 GTGTGTAACAAGCTTTTCGGGCCCGGGTAAAGAGTGGATGGCTTTATGGAAGTTAACTTA 4072  
 Qy 1304 LeuSerGlyPheMetValProSerGluAlaIleSerLeuSerGluThrValLysVal 1323  
 Db 4073 TTAAAGTGGCTTTATGAGTGGCTTCAAGAACAAATTTCTCTGACGACAGAGTGAAGAAAGTG 4132  
 Qy 1324 GluTyrAspHisGlyLysLeuAsnLeuTyrLeuAspSerValAsnGlnThrGlnPheCys 1343  
 Db 4133 GAATATGATCATGAAACTCAACCTTATTAGATTCTGTAAATGAAACCCAGTTTGT 4192  
 Qy 1344 ValAsnIleProAlaValAlaArgAsnPheLysValSerAsnThrGlnAspAlaSerValSer 1363  
 Db 4193 GTTATATTCCTCTCTGAGAAACCTTAAAGTTTCAAAATACCAAGATGCTTCAGTGTCC 4252  
 Qy 1364 IleValAspTyrTyrGluProArgArgGlnAlaValArgSerTyrAsnSerGluValLys 1383  
 Db 4253 ATAGTGATTAATCATAGCCCAAGAGACAGCGGTGAGAAATTAACAATCTTGAAGTGCAAG 4312  
 Qy 1384 LeuSerSerCysAspLeuCysSerAspValGlnGlyCysArgProCysGluAspGlyAla 1403  
 Db 4313 CTGTCTCTCGTACCTTGTGAGTGAATGTCAGAGGCTGCCGTCTTGTGAGATGAGAGCT 4372  
 Qy 1404 SerGlySerHisHisSerSerValIlePheIlePheCysPheLysLeuLeuTyrPhe 1423  
 Db 4373 TCAGGCTCCCATATCATCATCTTCAGTCATTTTATTTGTTTCAAGCTTCGTACTTT 4432  
 Qy 1424 MetGluLeuThrLeu 1428  
 Db 4433 ATGGAATTTGGCTG 4447  
 Qy Db 4433 ATGGAATTTGGCTG 4447  
 RESULT 7  
 AX534940 LOCUS 5895 bp DNA linear PAT 22-NOV-2002  
 AX534940 DEFINITION Sequence 7 from Patent WO02070696.  
 AX534940 ACCESSION  
 AX534940.1 GI:25261486  
 SOURCE  
 Homo sapiens (human)  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE  
 1 Schuh, A. and Sutherland, R.D.  
 Cdl09 nucleic acid molecules polypeptides and methods of use  
 Patent: WO 02070696-A 7 12-SEP-2002;  
 Schuh, Andre (CA) ; Sutherland, Robert D. (CA)  
 FEATURES  
 Location/Qualifiers  
 1..5895  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 113..4450  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAD57234.1"  
 /db\_xref="GI:25261487"  
 /translation="MGGPPLTPAHLVCYCTAALAVAPGRPLVTAPGIIIRPGNVTI  
 GVELLEHCPSQVTVKALKLTASNLTVSLAEAGVEKGSFRTLTLPLNSADBIY

ELRVTRTODELLFNSNTRLSFETKRIISVFIOQDKALYKPKQEVKRIYVLFSDPKPY  
KTSNLILIDPKSNLIQOWLSSQSDLGIVSKTFOLESHPILGMSIQOVNDQTYOS  
FOYSEYVLPKFEVTLQOTPLXGSMNSKHLGTLTAKTYKPYGVDVTLFPLSPFMGK  
KNKITKTFKINGSANSFENDEKKNVMSNGLSYELDSSPPGVILLTPTESVGI  
SNVSTNVFPEKHDIIEFPDYTVLKPSLNPATYKTRPAGNOCTLEBRNNVIT  
VYORATTEWSSGNSGNOMEAVOKINNTYTPOSGTYKIERPILDESSEIQKAYPLGS  
KSMASVHLSPKSPKTYIQKTRDENIKVGSPELVASGNKIKELSVNVRGQIYA  
VGNKONSTMSLSPENSWTPKACVIVYIISDDGEIISDVYKIPQVLFKXKIKLYMSKV  
KAPSEKVSILRISVTPDSIVGIVADVKNVIMANASNDIMENVAHELBYNNGYLG  
MFNSPFAVFOEGCLAVLTDLNLTQVYIDGVDNABYAEKPEMNEBHIYDHFSGS  
SPHVRKHPEPWTMLDNNMSRYOGEFETVPSITSWATGCVISEDGLGTTTPV  
ELQAFOPFTPLNLPYSVIRGEFFALEITTFNLTAKDITEVKITKSDKFDLIMTSE  
INATGHOQTLVPSVEDATVLEFIRPTHLGEIPIVTVALSPASDAVMTMIIWKAGI  
EKSYQSQIILDLTDNRLOSTLKTLSFSPBNVTGSERYQTALIGVLSPINGLASL  
IRMPYCGEONMINPANIYIIDLITFKKQOLTDNLKELKLSFMROGYSORELXYOREJ  
SFEAFGNVDPGSGTWSLAFVLRCELEADYIIDIQNVLHRTYTWLKGKXSGNEFDP  
GRVTHSELQGNKSPYTLNAYITYSLSLGRKTOPNDIVDSIHFLSESRSKISDNT  
LALITVALSSVGSPKAKELANMLTWBAEDGQMFWVSSEKLSDSQWRSKIDIEVAA  
YALSHFLOFQTSSEGIPIRMWLSRONSLSGGFASQDQTVALKALSEFALMNTERTN  
IQVTVGPSPSPVKFLIDTHNRLLOQTAELAVCPMAVNIISANGFPAICOLNVVN  
VKASGSSRRRSRISIONQEAFLDVAVKENEDDINHVDLNTCTSPSGRSGMALMEVNI  
LSGFWPSEALISLSTETKXKVEYDHGKLNILYDSVNETOCVNIPIARNKNTQDAS  
VSIYDIYPRRQAVRSYNSBVKLSCDLSDVQGCNRPCEBGASGSHHSSVIFICFK  
LITFMEML"

BASE COUNT 1708 a 1174 c 1241 g 1772 t  
ORIGIN

## Alignment Scores:

Pred. No.: 0 Length: 5895  
Score: 7317.50 Matches: 1426  
Percent Similarity: 98.75% Conservative: 1  
Best Local Similarity: 98.69% Mismatches: 1  
Query Match: 99.58% Indels: 17  
DB: 6 Gaps: 1

US-10-020-095-4 (1-1428) x AXS34940 (1-5895)

QY 1 MetGlnGlyProProLeuLeuThraAlaAlaHleuLeuCySerValCyThraAlaAlaLeu 20  
DB 113 ATGCAAGGCCACCGCTCTCGACCGCGCCCACTCTCGCTGCTGCGACCGCGCGCG 172  
QY 21 AlaValAlaProGlyProProPheLeuValThraAlaProGlyIleIleArgProGly 40  
DB 173 GCCGTGCTCCCGCGCTCTGCTTCTGCTGACGCCCGAGGATCATCGCGCGAGGA 232  
QY 41 AsnValThraIleGlyValGluLeuLeuGluHisCySerProSerGlnValThraValIleVal 60  
DB 233 AATGTACTAATGGGGTGGAGCTTCTGGAACTGCTGCTTCAACAGGTGACTGTGAAGGGG 292  
QY 61 GluLeuLeuYsThraIleAsnLeuThraValSerValIleuGluIleGluGlyValIlePhe 80  
DB 293 GAGCTCTCAAGACAGATCAACCTCACTGCTCTGCTGAGAGCAAGAGAGCTTT 352  
QY 81 GluYsGlySerPheYsThraLeuThraLeuProSerLeuProLeuAsnSerAlaAspGlu 100  
DB 353 GAAAAAGGCTCTTTAAGACACTTACTCTTCACTCACTGTAACAGTGCAGATGAG 412  
QY 101 IleTyGluLeuArgValThraGlyArgThraAspGluIleLeuPheSerAsnSerThra 120  
DB 413 ATTATGAGCTACGTGAACCGGACGTACCCAGAGAGAGATTTATCTCTAATAGTACC 472  
QY 121 ArgLeuSerPheGluThraYsArgIleSerValPheIleGluThraAspYsAlaLeuTy 140  
DB 473 CGCTATCATTTAGACCAAGAAATATCTGCTTCACTTCAACCAAGCAAGGCTTATAC 532  
QY 141 YsProYsGlnIleValYsPheArgIleValThraLeuPheSerAspPheYsProTy 160  
DB 533 AAGCCAAAGCAAGAGAGATTTCGATGTTACACTCTTCCAGATTTTAAAGCTTAC 592  
QY 161 YsThraSerLeuAsnIleLeuIleYsAspProYsSerAsnLeuIleGlnIleThraPleu 180  
DB 593 AAAACCTCTTAAACATTCATTAAAGACCCCAAAATTCAAATTTGATCCAAAGTGGTGG 652

QY 181 SerGlnGlnSerAspLeuGluValIleSerIyThraPheGlnLeuSerSerHisProIle 200  
DB 653 TCACAAACAAAGTATCTTGAGTGCATTTCCAAAACCTTTCCAGCTATCTTCCATCCAAATA 712  
QY 201 LeuGluAspTrpSerIleGlnValGlnValAsnAspGlnThraTyTyGlnSerPheGln 220  
DB 713 CTTGGGATCGGTCAATTCACATTCACAGGAATGACACAGACATTAATCAATTCATTCAG 772  
QY 221 ValSerGluTyValIleuProYsPheGluValThraLeuGlnThraProLeuTyCySer 240  
DB 773 GTTTACAAATATGATTTTACCAAAATTTGAGTGACTTTGCACACACCATTAATATGTTCT 832  
QY 241 MetAsnSerIyHisLeuAsnGlyThraIleThraAlaYsTyThraTyGlyYsProVal 260  
DB 833 ATGAATTCACATTTAAATGTAACCATCACGGCAAGTATACATATGGAAGCACAG 892  
QY 261 YsGlyAspValThraLeuThraPheLeuProLeuSerPheTyGlyYsYsYsAsnIle 280  
DB 893 AAAGGAGACGTAAACCTTACATTTTACCTTTATCTTTGGGAAAGAAATAATAT 952  
QY 281 ThraYsThraPheYsIleAsnGlySerAlaAsnPheSerPheAsnAspGluGluMetYs 300  
DB 953 ACAAAAACATTTAAAGATTAATGATCTGCAAACTTCTTTAAATGATGAAGAGTAA 1012  
QY 301 AsnValMetAspSerSerAsnGlyLeuSerGluTyYsLeuAspLeuSerSerProGlyPro 320  
DB 1013 AATGTATATGATTTCTTCAATATGACCTTTCTGTAATCTCGAGCTATCTTCCCTGACCA 1072  
QY 321 ValGluIleLeuThraThraValThraGluSerValThraGlyIleSerArgAsnValSerThra 340  
DB 1073 GTAGAAATTTTAAACACAGTGCACAAATCAGTAAACGATTAATCAAGAAATTAAGCACT 1132  
QY 341 AsnValPhePheYsGlnHisAspTyIleIleGluPhePheAspTyThraThraValLeu 360  
DB 1133 AATGTGCTTCAACCAACATATGATTAACATCATTTGATTTTGAATTAATCATCTGCTTG 1192  
QY 361 YsProSerLeuAsnPheThraIleThraValYsValThraArgAlaAspGlyAsnGlnLeu 380  
DB 1193 AAGCATCTCTCAACTTCAACGCCACTGGAAGGTACCTGCTGATGACCAACCACTG 1252  
QY 381 ThraLeuGluGluArgArgAsnAsnValIleThraValThraGlnArgAsnTyThraGlu 400  
DB 1253 ACTCTTGAAGAAAGAAATATATGATGATCAATACAGTACACAGAGAAATATACTGAG 1312  
QY 401 TyTrpSerGlySerAsnSerGlyAsnGlnYsMetGluAlaValGlnYsIleAsnTy 420  
DB 1313 TACTGAGCGGATCTTAACAGTGAATCAAGAAATGGAAGCTGTTCAAGAAATTAATAT 1372  
QY 421 ThraAlProGlnSerGlyThraPheYsIleGluPheProIleLeuGluYsAspSerGlu 440  
DB 1373 ACTGTCGCCCAAGTGAACCTTTAAGATTGAATTCCAATCTTGAAGATTCAGTGAG 1432  
QY 441 LeuGlnLeuYsAlaTyPheLeuGlySerIySerSerMetAlaValHisSerLeuPhe 460  
DB 1433 CTACAGTTGAAGGCTATTTCTTGTGTGTAAGAGATGAGCATGCTGCTT 1492  
QY 461 YsSerProSerIyThraTyIleGlnLeuYsThraGAspGluAsnIleYsValGly 480  
DB 1493 AAGTCTCTCAATGAACATCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 1552  
QY 481 SerProPheGluLeuValYsSerGlyAsnYsAspLeuYsGluLeuSerTyMetVal 500  
DB 1553 TCGCTTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1612  
QY 501 ValSerArgIyGlnLeuValAlaValGlyYsGlnAsnSerThraPheSerLeuThra 520  
DB 1613 GTATCAGGAGGACAGTGGTGGCTGTAGGAAACAAATTCACAAATGCTCTTTAACA 1672  
QY 521 ProGluAsnSerTrpThraProYsAlaCySerValIleValTyTyTyIleGluAspAspGly 540  
DB 1673 CCGAATAATCTTGAATCCCAAAAGCTGTGTAAATGATTAATTAATGAATATATGAG 1732  
QY 541 GluIleIleSerAspValLeuYsIleProValGlnLeuValPheYsAsnYsIleYs 560



Db 1733 GAAATTATTAAGTGTCTTAAAAATTCCTGTGCTGTTGTTTAAAAATAATAAG 1792  
 Qy LeuTyrTPSerLysValIleAlaGluProSerGluLysValSerLeuArgIleSerVal 580  
 Db 1793 CTATATTGGATTAAGTAAAGCTGAAACCATCTGAGAAAGCTCTCTTGGATCTCTG 1852  
 Qy ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600  
 Db 1853 ACACAGCCTGACCTCATGTGTTGGGATTTGAGTGTGCAAAAGTGTAATCTGATGAAT 1912  
 Qy AlaSerAsnAspIleThrMetGluAsnValIleHisGluLeuGluLeuTyrAsnThrGly 620  
 Db 1913 GCCCTTAATGATATTACATGATAAATGTGTGTCATGATGTGGAACCTTTAATACAGGA 1972  
 Qy TyrTyrLeuGluMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTyrVal 640  
 Db 1973 TATATTATGGCAGTCTCATGAAATCTTTGCACTCTTCAGGAATGTGACTCTGGGTA 2032  
 Qy LeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyrAspAsnAlaGluTyr 660  
 Db 2033 TTGACAGATGCAAACTCAGCAAGGATTAATGATGTTGTTTATGCAATGCAGAATAT 2092  
 Qy AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680  
 Db 2093 GCTAGAGGTTTATGAGAGAAATGAAAGACATATTGATGATATTCATGACTTTTCTTTG 2152  
 Qy GlySerSerProHisValArgLysHisAspPheGluThrTyrIleTPLeuAspThrAsn 700  
 Db 2153 GGTACAGCTCCACATGCTCGAAGACATTTTCAGAGACTTGATTTGGCTTACAGACCAAC 2212  
 Qy MetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIleThrSerTyr 720  
 Db 2213 ATGGGATCCAGATTTACCAAGAAATTTGAGTAATCTGACTGATTCATCATCTTCTTG 2272  
 Qy ValAlaThrArgLysPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrProVal 740  
 Db 2273 GTGGCTACTGTTTGTGATCTCTGAGAGACCTGGGCTTGGACTTACCAACTACCCAGTG 2332  
 Qy GluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArg 760  
 Db 2333 GAGCTCCAGCCTTCCAAACCACTTTTCTATTTTGTGATCTTCCCTACTCTGTTATCGA 2392  
 Qy GlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrGluVal 780  
 Db 2393 GGTGAGAAATTTGCTTGGAAATTAATCATATTATTTGAAAGATGCCAGAGGTT 2452  
 Qy LysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsn 800  
 Db 2453 AAGGTAAATCATTTGAGAAAGTGAACAATTTGATATTCTTAATGACTTCAAAATGAATAAAT 2512  
 Qy AlaThrGlyHisGlnGlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPhe 820  
 Db 2513 GCCACAGCCACACAGCAACCTTCTGTGCTCCAGTGAAGATGGGCACTGTTCTTTT 2572  
 Qy ProIleArgProThrHisLeuGlyGluIleProIleThrValIleThrAlaLeuSerProThr 840  
 Db 2573 CCCATCAGGCCAACACATCTGGGAGAAATTTCTATCAGTCAACAGCTCTTTCACCACT 2632  
 Qy AlaSerAspAlaValThrGluMetIleLeuValLysAlaGluGlyIleGluLysSerTyr 860  
 Db 2633 GCTTCTGATGCTGCACCCCAATGATTTTAAAGGCTGAAGAAATGAAGAAATTCATAT 2692  
 Qy SerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeu 880  
 Db 2693 TCACAAATTCATTTATTAAGATTTGATGACAAATAGGCTACAGATACCCCTGAAAACTTTG 2752  
 Qy SerPheSerPheProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIle 900  
 Db 2753 AGTTTCATATTCTCTCTTAATACAGTACCTGGCAGTGAAGATTCAGATACCTGCATTT 2812  
 Qy GlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGly 920

Db 2813 GGAAGATGTTCTTGCTCCCTCCATCAATGGCTTAGCCTCATGATTTGCGATCCTTATGGC 2872  
 Qy CysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThr 940  
 Db 2873 TGTGGTGAACAGAAACATGATTAATTTTGTCCCAATATTATTTACATTTTGGATTATCGACT 2932  
 Qy LysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPheMetArgGlnGly 960  
 Db 2933 AAAAAGAAACATCGACAGATTAATTTGAAAGAAAAGCTCTTTCATTTATGAGGCAAGGT 2992  
 Qy TyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyr 980  
 Db 2993 TACCGAGAAACCTTCTCTATCAGAGGAGATGGCTCTTTCAGTGTCTTTGGGAATTAAT 3052  
 Qy AspProSerGlySerThrTyrPheSerAlaPheValLeuArgCysPheLeuGluAlaAsp 1000  
 Db 3053 GACCTTCTGGGAGACCTTGTTGTGACGTTTGTGTTTAAGATGTTTCCCTGGAAGCGAT 3112  
 Qy ProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTyrLeuLysGlyHis 1020  
 Db 3113 CTTTACATGATATTGATCAGAAATGTGTTACACAGAAACATACACTTGGCTTAAAGACAT 3172  
 Qy GlnLysSerAsnGlyGluPheTyrAspProGlyArgValIleHisSerGluLeuGlnGly 1040  
 Db 3173 CAGAAATCCAAACGCTGAATTTTGGATCCAGAGAGTGATTTATGATGAGCTTCAAAGT 3232  
 Qy GlyAsnLysSerProValThrLeuThrAlaTyrIleValIleThrSerLeuLeuGlyTyrArg 1060  
 Db 3233 GCGCAATTAAGTCCAGTAACATTAACAGCTTAATATTGTAATCTTCTCCGGGATATAGA 3292  
 Qy LysTyrGlnProAsnIleAspValGlnGluSerIleHisPheLeuGluSerGluPheSer 1080  
 Db 3293 AAGTATCAGGCTTAACATTTGATGCAAGAGTCAATCATTTTGGAGTCAATTCAGT 3352  
 Qy ArgGlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGly 1100  
 Db 1101 SerProLysAlaLysGluAlaLeuAsnMetLeuThrTyrArgAlaGluGlnGluGly 1120  
 Db 3413 AGTCTTAACCGAAGAGAGCTTTGAATATGCTGACTTGGAGACAAAGAGAGTGGC 3472  
 Qy MetGlnPheTyrValSerSerGluSerLysLeuSerAspSerTyrGlnProArgSerLeu 1140  
 Db 3473 ATCAATTTCTGGGTGCATCAGAGTCCAAACTTTCGACTCCCGGACGACCTCCCTG 3532  
 Qy AspIleGluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGlu 1160  
 Db 3533 GATATTGAAGTGCAGCCTATGCACTGCTTCACACTTCTTACAAATTTTCAGACTTCTGAG 3592  
 Qy GlyIleProIleMetArgTyrPheSerArgGlnArgAsnSerLeuGlyPheAlaSer 1180  
 Db 3593 GGAATCCCAATTTATGAGTGTGCTGACAGCAAGAAATAGCTTGGGTGGTTTGCATCT 3652  
 Qy ThrGlnAspThrThrValAlaLeuLysValAlaLeuSerGluPheAlaAlaLeuMetAsnThr 1200  
 Db 3653 ACTCAGAGATCCACTGTGCTTAAAGGCTCTGTGCAATTTGCAACCCCTTAATGAATACA 3712  
 Qy GluArgThrAsnIleGlnValThrValThrGlyProSerSerProSerPro----- 1217  
 Db 3713 GAAAGGACAAATATCCAAAGTACCGTGAAGGCGCTTACCAAGTCTCTGTAAGTTT 3772  
 Qy -----LeuAlaValAlaGlnPro 1223  
 Db 3773 CTGATTGACACACAAACGCTTACTCTTCAGACACAGAGCTTGTGTGTGATCAGCCA 3832  
 Qy MetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValVal 1243  
 Db 3833 ATGGCAGTTAATATTTCCGGAATGTGTTTGAATTTGCAATTTGCAAGCTCAATGTGTA 3892  
 Qy TyrAsnValIleAlaSerGlySerSerArgArgArgSerIleGlnAsnGlnGluAla 1263  
 Db 3893 TATATGTGAAGGCTTCTGTGGTCTTCTAGAAAGCAGAAATCTATCCAAATCAAGAGCC 3952



QY 1264 PheAspLeuAspValAlaValIysGluAsnIysAspAspLeuAsnHisValAspLeuAsn 1283  
DB 3953 TTGATTTAGATGCTGCTGTAAGAAAGAAATGAAGATCTCATGTGAGTTTGAAAT 4012  
QY 1284 ValCysThrSerHisSerGlyProGlyIAspSerGlyMetAlaLeuMetGluValAsnLeu 1303  
DB 4013 GTGTGACAAAGCTTTTCGGGCGCGGTAGGAGTGCATGCTCTTATGGAAGTTAACTTA 4072  
QY 1304 LeuSerGlyPheMetValProSerGluAlaIleSerLeuSerGlyThrValIysIysVal 1323  
DB 4073 TTAAAGTGGCTTTATGGTGCCTTCAGAACCAATTTCTTGAGCAGACAGTGAAGAAAGTG 4132  
QY 1324 GluIlyAspHisGlyIysLeuAsnLeuIlyLeuAspSerValAsnGluThrGlnPheCys 1343  
DB 4133 GAATATGATCATGGAAGAACTCAACCTCTATTTAGATTCGTGAATGAAGAACCCAGTTTGTG 4192  
QY 1344 ValAsnIleProAlaValAlaGluAsnPheIysValSerThrGlnAspAlaSerValSer 1363  
DB 4193 GTTAAATATTCCTGCTGTGAGAACTTAAAGTTTCAATCCCAAGATGCTTCAGTGTCC 4252  
QY 1364 IleValAspIlyThrGluProAspGluAlaValArgSerIlyAsnSerGluValIys 1383  
DB 4253 ATATGATGATTACTATAGCCAGAGACAGCGCGGTAGAAATTAACAATCTGAGAGTGAAG 4312  
QY 1384 LeuSerSerCysAspLeuCysSerAspValGlnGlyCysArgProCysGluAspGlyAla 1403  
DB 4313 CTGTCTCTGCTGACCTTCGTCAGTGAATGTCAGGCTGCGCTCTGTGAGAGTGAAGCT 4372  
QY 1404 SerGlySerHisHisSerSerValIlePheIlePheCysPheIysLeuIlyThrPhe 1423  
DB 4373 TCAGGCTCCCATCATCATCTTCAGTCAATTTATTTTCTGTTCAAGCTTCTGTACTTT 4432  
QY 1424 MetGluLeuThrLeu 1428  
DB 4433 ATGGAACCTTGGCTG 4447  
RESULT 8  
AX534942 3535 bp DNA linear PAR 22-NOV-2002  
LOCUS AX534942  
DEFINITION Sequence 9 from Patent WO02070696.  
VERSION AX534942  
KEYWORDS AX534942.1 GI:25261490  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Bukarya; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1  
AUTHORS Schub, A. and Sutherland, R.D.  
JOURNAL Cdoi9 nucleic acid molecules polypeptides and methods of use  
Schub, Andre (CA); Sutherland, Robert D. (CA)  
Location/Qualifiers  
1. 3535  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
113. 3316  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD57235.1"  
/db\_xref="GI:25261491"  
CDS  
/translation="MGGPLTRAHLLCVCTAALAVAPGPRPLVTAIPGIIRGNVTI  
GVELLEHCPQVTVKALIKTASNLTVSLVLAEGVEKSPKTLPLSLNSADEIY  
ELAVRTQDEILFNSSTRISPEIKRISVFIQDKALYRKQEKVRIYTLSPDEKPY  
KTSILNLIKDKPSNLIQOMLSQSDGLVSKTFLSHSLIDGMSIQOVNODPYKYS  
FOYSEYVLPEKFEVTLQPLYCSMNSKHLNGITAKTYTKPVKGVDTLFLPSFMKG  
KNKITPEKINSANSFNDSEKONMDSNGISEYLDSSRPVPIITLYVESVGI  
SRVSTNVFPGKHDIIFEPDYTVLKPSLFTATKVTIRADSDNLTLEERRNVIT  
VTRNRYEWSGNSGNQKMEAVQKINLYPVSQGTGKPILEDSSSLQLAAYFVS  
KSMVAHSLFKSPSKTYIQIKTREDENIKVSGPELVSGNKRLEKELSNVAVSRGLVA  
VGKONSTMSLTPENSGWTPKACVIVYIEDGBEIIISDVAKIPQVLVFKKIKLYMSKV

KAEPSEKYSLRISVTQPDISIVGIVADKSNVLMNASNDITMENVHELELYNTGYLG  
MEMNFAFVQECGLAVLTDMANLTQYIGGVNDNAEYARFMEENGHIVDIDFELGS  
SPVAKHPEFTWIMLDMGYRIYQEFETVPDSTSWVAGFVISEDLGLTTPV  
ELIAROPPEIPLNIPYSVIRGEERFALETTINYLKDAIEVUJIEKSPFDIIMTNS  
INATROQTLVPSRSDGATVLFPIRPTHIGEIPITVTAISPTASDAVTOMILYKAGI  
EKSYISQSLILDDTDNRLOSITLKTSSFSPPPTVTSSEVQITAGIDVPSINGLASI  
IRMPYGCENQMINIPANPIYILDYLTKKOILDNDIKELASFMRGYRELLYQEDG  
SFSAFGNDVPSGTSLSAFVLRCPLEADPYIDIDQNVILHRTYTWLKHQKSGNEFWDP  
GRVISELQGGKSPVTLTAVYVTSILGVRKQYVERI"  
BASE COUNT 1108 a 675 c 731 g 1021 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 0 Length: 3535  
Score: 5470.00 Matches: 1062  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.91% Mismatches: 0  
Query Match: 74.44% Indels: 0  
DB: 6 Gaps: 0  
US-10-020-095-4 (1-1428) x AX534942 (1-3535)  
QY 1 MetGlnGlyProProLeuLeuThrAlaAlaHisLeuLeuCysValCysThrAlaAlaLeu 20  
DB 113 ATGCAAGGGCCACCGCTCCTGACCCGCCCACTCTGCTGCTGCTGCAACCGCGCTG 172  
QY 21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGlyGly 40  
DB 173 GCGGTGGCTCCCGGCGCTGCTGCTGAGAGCCCAAGGATATATACAGCGCGAGGA 232  
QY 41 AsnValThrIleGlyValGluLeuLeuGluHisCysProSerGlnValThrValIysAla 60  
DB 233 AATGTGACTATGGGGGTGAGCTTCTGGAACACTGCCCTTCAACAGTACTGTGAAGCG 292  
QY 61 GluLeuLeuIlyThrAlaSerAsnLeuThrValSerValLeuGluAlaGluGlyValPhe 80  
DB 293 GAGCTGCTCAAGACAGCATCAAACTCACTGCTCTGCTGGAAGCAGAAAGAGTCTTT 352  
QY 81 GluIlyGlySerPheIlyThrIleuThrLeuProSerIleuProLeuAsnSerAlaAspGlu 100  
DB 353 GAAAAGGCTCTTTAAAGACACTTACTCTTCATACATCACTTGAACAGTGCAGATGAG 412  
QY 101 IleIlyGluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThr 120  
DB 413 ATTATGAGCTACGCTTAACCGGACGTACCGAGATGAGATTTTATCTCTATATGATAC 472  
QY 121 ArgLeuSerPheGluThrIlyAspGlnIleSerValPheIleGlnThrAspIysAlaLeuIly 140  
DB 473 CGCTTATCATTTGAGACCAAGAGATATCTGCTTCATTTCAAAACAGACAGGCTTATAC 532  
QY 141 IlySerIlyGlnGluValIlyPheArgIleValIThrLeuPheSerAspPheIlyProIly 160  
DB 533 AAGCCAAAGCAAGAGTGAAGTTGTCATGTGTAACCTTCTCGATTTTAAAGCTTATAC 592  
QY 161 IlyThrSerLeuAsnIleLeuIleIlyAspProIlySerAsnLeuIleGlnGlnThrLeu 180  
DB 593 AAAACCTCTTTAAACATTTCTATTAGAAGACCCCAATCAAAATTTATTCACAGTGTG 652  
QY 181 SerGlnGlnSerAspLeuGlyValIleSerIlyThrPheGlnLeuSerSerHisAspProIle 200  
DB 653 TCACAAACAAAGATGATCTGAGATCATTTCCAAAACCTTTTACAGCTATCTTCCATCAATA 712  
QY 201 LeuGlyAspIlyPheIleGlnValGlnValAsnAspGlnThrIlyThrGlnSerPheGln 220  
DB 713 CTGTGATGCTGCTCTATTCAAGTCAAGTGAATGACCAACATATTATCAACATTTTACG 772  
QY 221 ValSerGluIlyValIleuProIlyPheGluValThrLeuGlnThrProLeuIlyCysSer 240  
DB 773 GTTTAGATATATGATTTACAAAATTTGAAGTGAATTTGACAGACCACTTATATTTGTTCT 832  
QY 241 MetAsnSerIlyHisLeuAsnGlyThrIleThrAlaIlyIlyThrIlyGlyIlyAspVal 260  
DB 833 ATGAATTTCTAAGCATTTAAATGATGATCAACGGCAAGTATACATATGGGAAGCCAGTG 892

Qy 261 LysGlyAspValThrLeuThrPheLeuProLeuSerPheTrpGlyLysValAsnIle 280  
Db 893 AAAGAGACGTAAACCTTACATTTTACCTTATCCTTTGGGAAAGAAATAATTT 952  
Qy 281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluMetLys 300  
Db 953 ACAAAACATTTAAGATTAATGAGCTGCAAACTTCTTTAATGATGAAGACATGAAA 1012  
Qy 301 AsnValMetAspSerSerAsnGlyLeuSerGlyValLeuAspLeuSerSerProGlyPro 320  
Db 1013 AATGTAATGAAATCTTCAATGAGACTTCTGAATACCTGGAATCTATCTTCCCTGGACCA 1072  
Qy 321 ValGluLeuLeuThrThrValThrGluSerValThrGlyIleSerArgAsnValSerThr 340  
Db 1073 GTAGAAATTTTAAACACAGTACAGAAATCACTTACAGTATTTCAAGAAAGTAAAGCACT 1132  
Qy 341 AsnValPhePheLysGlnHisAspTyrIleIleGluPhePheAspTyrThrThrValLeu 360  
Db 1133 AATGTCCTTCAAGCAACATGATTAATCATCTTGTGATTTTGAATTAATCACTGCTTG 1192  
Qy 361 LysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsnGluLeu 380  
Db 1193 AAGCATCTCTCACTTCAACAGCCACTGTGAAGGTAAGTCTGTGATGGCAACCACTG 1252  
Qy 381 ThrLeuGluGluArgAsnAsnValValIleThrValThrGluArgAsnTyrThrGlu 400  
Db 1253 ACTCTGAAGAAAGAAATTAATGTACTTAACAGTACAGACAGAAATTAATACAGAG 1312  
Qy 401 TyrTrpSerGlySerAsnSerGlyAsnGluLysMetGluAlaValGlnLysIleAsnTyr 420  
Db 1313 TACTGAGCGGATCTAAACAGTGAATTCAGAAATTCAGAACTGTTCAAGAAATTAATAT 1372  
Qy 421 ThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGluAspSerSerGlu 440  
Db 1373 ACTGCCCCCAAGGGAACCTTTAAGATTAATTCCAATCCCTGGAGATTCACAGTAG 1432  
Qy 441 LeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSerLeuPhe 460  
Db 1433 CTACAGTTGAAGGCTATTTCTTGGTAGTAAAGTAGCAAGTTCATAGTCTGTTT 1492  
Qy 461 LysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLysValGly 480  
Db 1493 AAGTCTCTAGTAAACATATCATCTCACTAAACCAAGACATGAATAATTAAGTGGGA 1552  
Qy 481 SerProPheGluLeuValValSerGlyAsnLysArgLeuLysGluLeuSerTyrMetVal 500  
Db 1553 TCGCCTTTTGAAGTGTGGTGTAGTGGCAAAACGATTTGAAGGATTAAGCTATATGTA 1612  
Qy 501 ValSerArgGlyGlnLeuValAlaValGlyLysGlnAsnSerThrMetPheSerLeuThr 520  
Db 1613 GTATTCAGGGGACAGTGTGGTGTAGGAAACAAATTCACAAATCTTCTTTAAACA 1672  
Qy 521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGluAspAspGly 540  
Db 1673 CCAGAAATTTCTTGACCTCCAAAGCCGTGTAAATGTGTATTAATGAAGATGATGGG 1732  
Qy 541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560  
Db 1733 GAAATTAATAGTGTCTTAAATAATTCCTGTGACGTTGTTTAAATAATTAATGAAG 1792  
Qy 561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580  
Db 1793 CTATATTTGAGTAAGTAAGTAAGTCAACCATCTGAGAAAGTCTCTCTTGAATCTCTGTG 1852  
Qy 581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600  
Db 1853 ACACAGCCTGACCTCATATGTTGGATTTAGCTGTGCAAAAGTGAATCTGATGAT 1912  
Qy 601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeuGluLeuTyrAsnThrGly 620  
Db 1913 GCCTCTATATGATATTACAAATGAATAATGTGTCCATGATGAGTGAATCTTATTAACAGAG 1972

Qy 621 TyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTrpVal 640  
Db 1973 TATTATTAGGAGTATTCATGAATCTTTTGCAGTCTTTCAAGAAATGTGACTCTGGTA 2032  
Qy 641 LeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyrAspAsnAlaGluTyr 660  
Db 2033 TTGACAGATCAAACTTCACAGAGAGATTAATTAATGAGTGTATTATGACAAATGACAAATAT 2092  
Qy 661 AlaGluArgPheMetGluGluAsnGluLysIleValAspIleHisAspPheSerLeu 680  
Db 2093 GCTGAGAGGTTTATGAGAGAAATGAAGACATATGTAGATATTCATGACTTTTCTTTG 2152  
Qy 681 GlySerSerProHisValArgLysHisPheProGluThrTrpIleTrpLeuAspThrAsn 700  
Db 2153 GTTAGCAGTCCACATGCCAAGACATTTTCCAGACATTTGGATTTGGCTAGACACCAAC 2212  
Qy 701 MetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIleThrSerTrp 720  
Db 2213 ATGGGTTACAGGATTTTACAAAGAAATTTGAAGTAACTGATCTGATCTATCAGCTTCTGG 2272  
Qy 721 ValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrProVal 740  
Db 2273 GTGGCTACTGGTTTGTGATCTCTGAGAGACTGGGTCTTGACTTAACACTACTCCAGTG 2332  
Qy 741 GluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArg 760  
Db 2333 GAGCTCAACGCTTCCAAACCATTTTTCATTTTGAATCTTCCCTACTCTGTTATCAGA 2392  
Qy 761 GlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrGluVal 780  
Db 2393 GGTGAAGAAATTTCTTGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2452  
Qy 781 LysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsn 800  
Db 2453 AAGGTATCTTATGAGAAATGACAAATTTGATTTCTAATGACTTCAATGAAATTAAT 2512  
Qy 801 AlaThrGlyHisGlnGlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPhe 820  
Db 2513 GCCACAGGCGCACAGACAGACCTTCTGGTCCACAGTAGAGTGGGCAACTGTTCTTTT 2572  
Qy 821 ProIleArgProThrHisLeuGlyGluIleProIleThrValThrAlaLeuSerProThr 840  
Db 2573 CCCATCAGGCCAACATCTGGAGAAATTCATATCAAGTCAACACTCTTCAACCACT 2632  
Qy 841 AlaSerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIleGluLysSerTyr 860  
Db 2633 GCTTCTGATGCTTCAACGATGATTTTATGTAAGGCTGAAGAAATGAATAATATAT 2692  
Qy 861 SerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeu 880  
Db 2693 TCACATTCATCTTATTAAGCTTGACCTGACCAATAGGCTACAGATGCCCTGAATACTTTG 2752  
Qy 881 SerPheSerPheProPheAsnThrValThrGlySerGluArgValGlnIleThrAlaIle 900  
Db 2753 AGTTTCTCATTTCTCTATATACAGTACTGGCAGAGAAAGATTCAGATCATCTGAAT 2812  
Qy 901 GlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGly 920  
Db 2813 GAGAGATTTCTTGTCTTCTTCAATCAATAGGCTTACCTCATTAATTTGATGATGCTTATGGC 2872  
Qy 921 CysGlyGlnGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThr 940  
Db 2873 TGTGTGAACAGAAACATGAATTAATTTTGTCCCAATATTTACATTTTGGATTATCGACT 2932  
Qy 941 LysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPheMetArgGlnGly 960  
Db 2933 AAAAAGAAACAACTGACAGATTAATTTGAAGAAAGAAAGCTTTCATTTATGAGCAAGT 2992  
Qy 961 TyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyr 980  
Db 2993 TACCAGAGAACTTCTATACAGAGAGAGATGAGCTTTTCAAGTCTTTTGGGAAATAT 3052  
Qy 981 AspProSerGlySerThrTrpLeuSerAlaPheValLeuArgCysPheLeuGluAlaAsp 1000

Db 3053 GACCCCTTGGAGGACCTTGGTTCAGCTTTTGAAGTGTTCCTTAAGCCGAT 3112  
Qy 1001 ProTyrlleAsp11eAsp11aenValleuH1sArgThrTyThrTrpleuLysGlyHs 1020  
Db 3113 CCTTACATGATATTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3172  
Qy 1021 GlnLysSerAsnGlyGlnLysPheAspProGlyArgVal11eHisSerGlnLeuGlnGly 1040  
Db 3173 CAGAAATCCACGCGTCAATTTTGGATCCAGGAAGATGATTCATGATGATGATGATGAT 3232  
Qy 1041 GlnLysSerProValThrLeuThrAlaTyrlleValThrSerLeuLeuGlyTyArg 1060  
Db 3233 GGCATTAATAAGTCCAGTACCTTACAGCTATATGTAACCTCTCTCCGAGATATAGA 3292  
Qy 1061 LysTyrlGln 1063  
Db 3293 AAGTATCAG 3301  
RESULT 9  
AX534944 3535 bp DNA linear PART 22-NOV-2002  
LOCUS AX534944  
DEFINITION Sequence 11 from Patent WO02070696.  
ACCESSION AX534944  
VERSION AX534944.1 GI:25261494  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE  
AUTHORS Schuh, A. and Sutherland, R.D.  
TITLE Cd109 nucleic acid molecules polypeptides and methods of use  
JOURNAL Patent: WO 02070696-A 11 12-SEP-2002:  
Schuh, Andre (CA) ; Sutherland, Robert D. (CA)  
FEATURES  
source  
1. .3535  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
113. .3316  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD57236.1"  
/db\_xref="GI:25261495"  
/translation="MQGPEPLTRAHLLCYCTAALAVAPRPLVTAAPGIRGCAVTL  
GVELLHCPQSQTVAELKLTASNLTVSLAEVFEKGSFTTLPLSLPLNSADELY  
ELRVGRQDELIFSNSTRLSPEKRI SVFIQTDKALYKQEVERIVTLPSDFKPY  
KTSNLILKIDPKSNLIQOWLSQOSDLGVSKTQPSHPIIDMSIQOVNDQTYOS  
FQVSEVILPKFEVTLQTPLYGSMNSKHLNGTTAKYTGKPKGVTLTLPILSPWGK  
KNITKTEPKINGSANPSFNDDEMKVMNSGLSEYLDLSGPEVILLTTVSTGTI  
SRVSTNVPFKQHDYIIEFDYTTVLKSLNFTATVKTADGNOLLEBRNNVIT  
VQRNVTEWSGNSNGNOMEAVOKINTVPOSGFKLEPILDESSSELQAYPLGS  
KSMNVHSLFKSPKTYIQKTRDENIKVSGPELVSGNKLEKELSTVNSRGLVA  
VKNSTMSLSLTPENSMTPKACVIVYIEDDEGISDYLKIPEVOLFKNKVIKLSYK  
KAPEKSKLSISTOPDSIVGIVAVDSVNMASNDITMENVVHLELVTGYIG  
MEMNSFAYROBGILVLDNANTKQYIGVYNMAYARERPMENGHVYDIDESLGS  
SEVRKHPFETIWTLDNNGKRIYGEFEVTPDSITSWATFVISEDGLATTPV  
ELQAFQPFITPLNLYSVIRGEPALEITFENYLADAEVKYIIEKSKFDLITPNS  
INATHQOQLLVSESDGATVLPRIPTHLEIPIVTALSPASADVOMILVKAGLI  
EKSYQSILDLIDNRLOSTLKTLSFSPPTVTSERVOITADIVGSPINGASLI  
IRMPYGCGRONMINFAPNITLYLDYLTKKKOLTDNLKEXKLSFMRGXYORELIXOREG  
SFSARGVDRGSTWLSAFLVLCPLADPDIIDONVLRHTYTLKSHOKSGERFMDP  
GRTVISELQGNKSPVTLNAYIVLSLGRKTYQFRFRI"

DB: 6 Gaps: 0  
US-10-020-095-4 (1-1428) x AX534944 (1-3535)  
Qy 1 MetGlnGlyProProLeuLeuThrAlaAlaHisLeuLeuCyValCysThrAlaAlaLeu 20  
Db 113 ATGAGGAGCCCAACCCCTCTGACCGCCGCCCACTCTCTGCTGTGTGACCGCCGCTG 172  
Qy 21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyLysLeuArgProGlyGly 40  
Db 173 GCCGTGCTCCCGGCTCCGCTGCTTCTGTGTGACAGCCCAAGGATATACAGCCCGCTG 232  
Qy 41 AsnValThrIleGlyValGlnLeuLeuGlnHisCysProSerGlnValThrValLysAla 60  
Db 223 AATGTACCTATTGGGGGTGAGCTTCTGGAACACTGCCCTTCAACAGTACGTGAGAGGG 292  
Qy 61 GlnLeuLeuLysThrAlaSerAsnLeuThrValSerValLeuGlnAlaGlnGlyValPhe 80  
Db 293 GAGCTGCTCAAGACAGCATCAAACTCACTGCTCTGCTGGAAGCAGAAAGAGCTCTT 352  
Qy 81 GlnLysGlySerPheLysThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGlu 100  
Db 353 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCTACTGTAACAGTGCAGATAG 412  
Qy 101 IleTyrlGlnLeuArgValThrGlyArgThrGlnAspGlnIleLeuPheSerAsnSerThr 120  
Db 413 ATTTATAGCTACGTCGTAAACCGAGCTACCCAGATGAGATTTTATCTCTAATGTAAC 472  
Qy 121 ArgLeuSerPheGlnThrLysArgGlyLeuSerValPheIleGlnThrAspLysAlaLeuTy 140  
Db 473 CGCTTATCATTTGAGACACAGAGATATCTGCTTCAATCAACAGACAGAGCCCTTATAC 532  
Qy 141 LysProLysGlnGlnValLysPheArgGlyValThrLeuPheSerAspPheLysProTy 160  
Db 533 AAGCCAAAGCAGAAAGTAAAGTTGCACTGTGTACCTCTTCCGATTTTAAAGCCTTATAC 592  
Qy 161 LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGlnThrLeu 180  
Db 593 AAAACCTCTTTAAACATCTCTATTAAGACCCCAATCAATTTGATCAACAGTGTG 652  
Qy 181 SerGlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIle 200  
Db 653 TCACAAACAAAGATCTTGAAGTCACTTCCAAAACCTTTACAGCTATCTCCCATCAATA 712  
Qy 201 LeuGlyAspTPSerIleGlnValGlnValAsnAspGlnThrTyrlGlnSerPheGln 220  
Db 713 CTGTGTGACGTGCTCATTTCAAGTTCAAGTGAATGACAGACATATTATCAATTCATTCAG 772  
Qy 221 ValSerGlnTyrlValLeuProLysPheGlnValThrLeuGlnThrProLeuTyrlCysSer 240  
Db 773 GTTTAGAAATGTAATTCACAAATTTAAGTACCTTTCAGACACCAATTATATTGTTCT 832  
Qy 241 MetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrlTyrlGlyLysProVal 260  
Db 833 ATGATTTCTTAAGCATTTAAATGATGATCAACAGGAAAGTATACATATGGGAAGCAGTG 892  
Qy 261 LysGlyAspValThrLeuThrPheLeuProLeuSerPheTrpGlyLysLysLysAsnIle 280  
Db 893 AAAGAGACGTAACCTTACATTTTACCTTTATCTTTGGGGGAAAAGAAAATAATTT 952  
Qy 281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGlnLysLys 300  
Db 953 ACAAAACATTTAAGATTAATGATTCGAAACTTCTTTTAAAGATGAAGAGATGA 1012  
Qy 301 AsnValMetAspSerSerAsnGlyLeuSerGlnTyrlLeuAspLeuSerSerProGlyPro 320  
Db 1013 AATGTATGATGATTTCTTCAATGAGACTTCTGTAATACCTGATCTATCTTCCCTGAGACA 1072  
Qy 321 ValGlnIleLeuThrValThrGlnLysValThrGlyIleSerArgAsnValSerThr 340  
Db 1073 GTAGAAATTTTAAACACAGTACAGAAACAGTATTCAGAAAGATGAAGCACT 1132  
Qy 341 AsnValPhePheLysGlnHisAspTyrlIleIleGlnPhePheAspTyrlThrValLeu 360

## Alignment Scores:

Pred. No.: 0  
Score: 5461.00  
Percent Similarity: 99.91%  
Best Local Similarity: 99.81%  
Query Match: 74.32%

Length: 3535  
Matches: 1061  
Conservative: 1  
Mismatch: 1  
Indels: 0

```

Db      1133 AATGTCCTTCAAGCAACATGATTCATCATGTTTGTGATTTACTACGTCCTGG 1192
Qy      361 LysProSerLeuAanPheThrAlaThrValLysValThrArgAlaAspGlyAsnGluLeu 380
Db      1193 AAGCATCTCTCAACTTCACAGCCACTGTGAAGGTAAGTCTGTCGATGGCAACCAACTG 1252
Qy      381 ThrLeuGluGluArgAsnAsnValValIleThrValThrGlnArgAsnTyrThrGlu 400
Db      1253 ACTCTTGAAGAAAGAAATATAGTAGTCATTAACAGTACACAGAGAAATATATCTAGAG 1312
Qy      401 TyrTrpSerGlySerAsnSerGlyAsnGlnLysMetGluAlaValGlnLysIleAsnTyr 420
Db      1313 TACTGAGCGGATCTCAACAGTGAATCAGAAATGCAAGCTGTTCCAGAAATTAATATAT 1372
Qy      421 ThrValProGlnSerGlyThrPheLysIleGlnPheProIleLeuGluAspSerSerGlu 440
Db      1373 ACTGCCCCCAAGGGAACCTTTTAAGATGAATTCCTCAATCCAGAGGATTCACAGTAGAG 1432
Qy      441 LeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSerLeuPhe 460
Db      1433 CTACAGTTGAAGGCTTATTTCTTGTAAGTAAAGTACATGGCAGTTCAATAGTCTGTTT 1492
Qy      461 LysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLysValGly 480
Db      1493 AAGTCTCTAGTAAACATACATCCAACTAAACCAAGAGATGAATAATATTAAGGTGGGA 1552
Qy      481 SerProPheGluLeuValValSerGlyAsnLysArgLeuLysGluLeuSerTyrMetVal 500
Db      1553 TCGCCTTTTGAGTGGTGTAGTAGGCAACCAACATTAAGAGATTAAAGCTATATAGTGA 1612
Qy      501 ValSerArgGlyGlnLeuValAlaValGlyLysGlnAsnSerThrMetPheSerLeuThr 520
Db      1613 GTATCCAGGGGACAGTTGGTGGCTGTAGAAAACAAATTCACAAATGTTCTTTTAAACA 1672
Qy      521 ProGluAsnSerTyrPheProLysValAcysValIleValTyrTyrIleGluAspAspGly 540
Db      1673 CCAGAAATTTCTTGACCTCCAAAGCCCTGTAAATGTGTATTAATTAAGAGATGATCGG 1732
Qy      541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560
Db      1733 GAATTAATTAAGTGAATGTTCTAAATAATTCCTGTACGCTGTTTAAATAATTAAGATAAG 1792
Qy      561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580
Db      1793 CTATATTTGAGATTAAGTAAAGCTGAACCATCTGAAGAGTCTCTTTAGATCTCTGTG 1852
Qy      581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600
Db      1853 ACACAGCCTGACCTCATATGTTGGGATTTGATCTGTGCAAAAGTGAATCTGATGAT 1912
Qy      601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeuGluLeuTyrAsnThrGly 620
Db      1913 GCCCTTAATGAATATTAACATGAATAATGTGTCCATGATGTGAACCTTTATTAACAGAGA 1972
Qy      621 TyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTyrVal 640
Db      1973 TATATTTTAGGCACTTCATATTAATTTCTTTGACGCTTTTCAGAAATGTGGACTCTGGGTA 2032
Qy      641 LeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyrAspAsnAlaGluTyr 660
Db      2033 TTGACAGATGCAAACTCAGCAAGGATTAATATGATGTGTTATGACAAATGACGAATAT 2092
Qy      661 AlaGluArgPheMetGluGluAsnGluLysIleValIleAspIleHisAspPheSerLeu 680
Db      2093 GCTAGAGAGTTTATGAGAGAAATGAAAGACATATTTGATGATATTCAGACTTTTCTTTG 2152
Qy      681 GlySerSerProHisValArgLysHisPheProGluThrTyrIleTyrLeuAspThrAsn 700
Db      2153 GGTAGCACTCCACATGTCGAAAGCATTTTCCAGAGACTTGGATTTGGCTTAGACCAAC 2212
Qy      701 MetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIleThrSerTyr 720

```

---

```

Db      2213 ATGGGTTCCAGAGTTTACCAAGAAATTTGAAGTAACGTACGTATCTATCACTTCTGG 2272
Qy      721 ValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrProVal 740
Db      2273 GTGGCTACTGGTTTGTGATCTCTGAGAGACCTGGGCTTGGACTATACACTACTCCAGTG 2332
Qy      741 GluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArg 760
Db      2333 GAGCTCCAACTCCCAACCATTTTTCATTTTGTGAATCTTCCCTACTCTGTTATCAGA 2392
Qy      761 GlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrGluVal 780
Db      2393 GGTGAATAATTTCTTGGAAATTAACATATTAATTAATTAATTAATTAATTAATTAATTA 2452
Qy      781 LysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsn 800
Db      2453 AAGGTATTCATTGAGAAAGTGAACAAATTTGATATTTATGACTTCAATGAAATTAAT 2512
Qy      801 AlaThrGlyHisGlnGlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPhe 820
Db      2513 GCCACAGGCCACAGCAGACCTTCTGTTCCACAGTAGAGATGGGCACTGTTCTTTT 2572
Qy      821 ProIleArgProThrHisLeuGlyGluIleProIleThrValThrAlaLeuSerProThr 840
Db      2573 CCATCAGGCCAACACATCTGGAGAAATTCCTATCACATTCACAGCTCTTTCAACCCT 2632
Qy      841 AlaSerAspAlaValThrGlnMetIleLeuValIleValGluGlyIleGluLysSerTyr 860
Db      2633 GCTTCTGATCTGTACCCAGATGATTTTGTAAAGCTTAAGAGAAATGAAATATCATAT 2692
Qy      861 SerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeu 880
Db      2693 TCACAAATCCATCTTATTAAGACTGACATGACAAATAGGCTACAGATCACTCGAAACCTTG 2752
Qy      881 SerPheSerPheProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIle 900
Db      2753 AGTTTCATATTTCTCTTAATACAGTACAGTGAAGAAAGTTCAGATCAGTCGAAT 2812
Qy      901 GlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGly 920
Db      2813 GGAATGTTCTTGCTCTTCCATCATGATGCTTACGCTTATGATGCTGATGCTTATGGC 2872
Qy      921 CysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThr 940
Db      2873 TGTGTGAACAGAAACATGAATATTTGCTCCAAATATTTACATTTGATATTCGACT 2932
Qy      941 LysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPheMetAlaGlnGly 960
Db      2933 AAAAAGAAACAACTGACAGATTAATTTGAAAGAAAGCTCTTTCATTTATGAGGCAAGGT 2992
Qy      961 TyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyr 980
Db      2993 TACCAGAGAAACCTCTCTATCAAGAGGAAAGAGGCTTTTCAGTGTGTTTGGAAATAT 3052
Qy      981 AspProSerGlySerThrTyrPheLeuSerAlaPheValLeuArgCysPheLeuGluAlaAsp 1000
Db      3053 GAOCCTTCTGGAGACACTTGTGTGTACGCTTTTGTTTAAGATGTTTCTTGAAGCCGAT 3112
Qy      1001 ProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTyrLeuLysGlyHis 1020
Db      3113 CCTTACATGAATATTTGATGAGAAATGTGTACACAGAAACATCACTTGGCTTAAAGGACAT 3172
Qy      1021 GluLysSerAsnGlyGluPheThrAspProGlyArgValIleHisSerGluLeuGlnGly 1040
Db      3173 CAGAAATCCAAAGGTGAATTTGGATCCAGAAAGTGAATTCATATGAGCTTCAAGGT 3232
Qy      1041 GlyAsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArg 1060
Db      3233 GGCATTAAGTCCAGTAACATTAACGCTATATTTACTTCTCTCCGGGATATAGA 3292
Qy      1061 LysTyrGln 1063
Db      3293 AAGTATCAG 3301

```

[illegible]

Db	1360	TCAGAAATGGCAGACGACCAATGCTGTTTTCAGCAACAATGATATCATCATTTGAGATTTT	1417
Qy	355	AspYrThrThrValLeuIleuysProSerIeuaenphenrIalathrValIysValThrArg	374
Db	1420	GATTATACGACTGCTTGAAACCATCTCTCAACTTCACAGCCACTGTGAAGGTGAGCCGC	1479
Qy	375	AlaAspGlyAsenGlnLeuThrLeuGlnGluIuArgaAsnValValIleThrValThr	394
Db	1480	TCTGATGGCAATCAACTAATCTCTAGAAATAAGAAATATGATTTGGTCAAGTATGACC	1539
Qy	395	GlnArg-----AsnYrThrGluYrTrpSerGlySerAsnSergIyAsnGlnIuMet	412
Db	1540	CAGCGGAAAAATACACCCTTGAG-----AGTCAGAGAGATCAGGAATG	1584
Qy	413	GlnAlaValGlnIuysIleAsnYrThrValProGlnSergIyThrPheIysIleGluPhe	432
Db	1585	GATTATATCCAGACCGTAATATATCTATCCCTTCCCTATTAAGATCAGTACATTC	1644
Qy	433	ProIleLeuGluAspSerSergIuLeuGlnIleuIysAlaYrPheLeuGlySerIysSer	452
Db	1645	CCAGCAAGTCCATTTCCGGCAGGCTACAACTGAAGGCTTATTTCTTGATGTACAGT	1704
Qy	453	SerMetAlaValHisSerLeuPheIysSerProSerIysThrYrIleGlnIleuYthr	472
Db	1705	TCTGGACAGCTCCACGATATGTCTCCTCTCTATTAAGATCAGTACATTCAGCTTAAACC	1764
Qy	473	ArgAspGluAsnIleIysValGlySerProPheGlnLeuValIleSergIyAsnIuysArg	492
Db	1765	AGAGATGAATATATTAAGGTGGGCTCACCTTTGATTTGATGTATAGTGGCAACAGACA	1824
Qy	493	LeuIysGlnLeuSerYrMetValIleSerArgGlyGlnIleuValAlaValGlyIysGln	512
Db	1825	TTCAAGGACTTAAGCTAATATGTATATCAAAAGGACGTTGGTGGCTGCAGGCAACAA	1884
Qy	513	AsnSerThrMetPheSerLeuThrProGlnAsnSerTrpThrProIysAlaCysValIle	532
Db	1885	AGTTCAGACCTTCTCTTTTACACACAGGCGTCTCTGGCTCCAAAGGCGCTGATATATC	1944
Qy	533	ValYrYrIleGluAspAspGlyGlnIleIleSerAspValLeuIysIleProValGln	552
Db	1945	GCATATTATATTCGACAGAGATGGGGAATTTAATGATATTTCTAAATATCCCGTTCAG	2004
Qy	553	LeuValPheIysAsnIuysIleIysLeuYrTrpSerIysValIysAlaGlnProSergIu	572
Db	2005	CTTGCTTTTGAATAATAGATTAAGCTGTTTGGAGTAACTACAGTCAACCATCTGAT	2064
Qy	573	LysValSerLeuArgIleSerValThrGlnProAspSerIleValGlyIleValAlaVal	592
Db	2065	AAGGCTCTCTCAGGATCTCTGGCAGACAGTCTGATCTCCCTGGGGGATGTGAGCTGT	2124
Qy	593	AspYrsSerValAsnLeuMetAsnAlaSerAsnAspIleThrMetGlnAsnValHis	612
Db	2125	GACAAAAGTGAACCTTATGAAAACTCCAAACAGTATTTCAATGGAACCATGTCAT	2184
Qy	613	GluIeuGlnLeuYrAsnThrGlyYrYrIeuGlnIuMetPheMetAsnSerPheAlaVal	632
Db	2185	GAACTGGAACCTTATATACACAGAAATATTTATCTAGGACATGTCATGAATCTCGCTGTA	2244
Qy	633	PheGlnGluCysGlyLeuTrpValLeuThrAspAlaAsnLeuThrIysAspYrTrlLeAsp	652
Db	2245	TTTCAGGAATGGGCTCTGGGTGTGACAGATGCMACTTATATAGATAGCATTCAT	2304
Qy	653	GlyValYrArgAsnAlaGluYrIuArgIuArgPheMetGlnIuAsnGlnGlyHisIle	672
Db	2305	GAACTCTACGATCTGAAGAGATTTCTAAACGCTTGCAAGAGAAATAGAGCAAACTG	2364
Qy	673	ValAspIleHisAspPheSerLeuIysSerProHisValArgIysHisPheProGln	692
Db	2365	GTTGATTTTGAAGAGCTTCTTCAAGTAAACAATGTACATGTACAGGAATAATTTTCCAGAA	2424
Qy	693	ThrTrpIleTrpLeuAspThrAsnMetGlyYrArgIleYrGlnGlnPheGluValThr	712
Db	2425	ACATGATTTTGGCTAGACGCTTACCGGATCTGAAGTATTTAGAAATTTGAAGTTTCT	2484

QY	713	ValProAspSerIleThrSerTyrValAlaThrGlyPheValIleSerGluAspLeuGly	732
Db	2485	GTACTCGACTCCATCACTCTTGGGTGGCTTGGCTTTTGTCATCTAGAAAGACTGGCT	2544
QY	733	LeuGlyLeuThrTrpTrpProValGlyLeuGlnAlaPheGlnProPheIlePheLeu	752
Db	2545	TTTGGGTTAACACCGCTTCCAGCAGAGCTGCACACCTTCCAAACCGTTTTCCTTCTG	2604
QY	753	AsnLeuProTyrSerValIleArgGlyGluGluPheAlaLeuGluIleThrIlePheAsn	772
Db	2605	AATCTTCGCTACTCTGTTATCAGAGGGTGAAGAGTTGCTTTGGAAAGATCATCTGCAAT	2664
QY	773	TyrLeuIysAspAlaIleTrpValIleValIleIleGluIysSerAspLysPheAspIle	792
Db	2665	TATTTGAAGAAGTACAAATTTAAAGTTGTATCTACTATTGAGGAAGGTGACAGCTTTGAATTT	2724
QY	793	LeuMetThrSerSerGluIleAspAlaThrGlyIleGlnGlnTrpLeuLeuValProSer	812
Db	2725	TTAATGACTTCAATATGACACCAATGGCACCATATATACGGAAGACTGTTCAGATTCCACG	2784
QY	813	GluAspGlyAlaIleTrpValIlePheProIleArgProThrIleLeuGlyGluIleProIle	832
Db	2785	GATATATGGGTCACACTCTGTTTCCCATCAAGCCACAGCACTTGGAGAGATTCCTATC	2844
QY	833	ThrValThrAlaLeuSerProThrAlaSerAspAlaValThrGlnMetIleLeuValLys	852
Db	2845	ACAGTGAAGCGGTGCTCACCACTGCTGTGATCTGTCCACCCAGACAATTTGTGTAAAG	2904
QY	853	AlaGluGlyIleGluIysSerTyrSerGlnSerIleLeuLeuAspLeuThrAspAsnArg	872
Db	2905	CCTGAAGGGAATGAAGAAGTGTATTCGAATCTGTCTTATTTGATGATCGACAGATAGCAAC	2964
QY	873	LeuGlnSerThrLeuLysThrLeuSerPheSerPheProProAsnThrValThrGlySer	892
Db	2965	GTAGAAGTAAAGCAGCAATCATATGAGTATCTCTTCTCTGTATAGGTATTTGGCAGT	3024
QY	893	GluArgValGlnIleThrAlaIleGlyAspValLeuGlyProSerIleAsnGlyLeuAla	912
Db	3025	GAAAGAGTTCCAGATCACAGCAATTTGGAGATATCTTGTTCTTCATCAATATGGCTTATCT	3084
QY	913	SerLeuIleArgMetProTyrGlyCysGlyGluGlnIleMetIleAsnPheAlaProAsn	932
Db	3085	TCAGTGAATCCGGAAGCCCTTATGATGGTGTAAACAGAAACATGATATTTTGTCTCCAAAT	3144
QY	933	IleTyrIleLeuAspTyrLeuThrLysIleGlySerGlnLeuThrAspAsnLeuLysGlyLys	952
Db	3145	ATTTCATCTTTGGATTATCTGACTTAACACAGAAACAGCTGACAGTATTTTAAAGAAAG	3204
QY	953	AlaLeuSerPheMetArgGlnGlyTyrGlnArgGluLeuLeuTyrGlnArgGluAspGly	972
Db	3205	GCCCTTCATATATGAGGCAAGGTTTACCAAGGAGCTTCTATCAGAGGAAAGATGGC	3264
QY	973	SerPheSerAlaPheGlyAsnTyrAspProSerGlySerThrTrpLeuSerAlaPheVal	992
Db	3265	TCCCTTCAGTGGCTTTGGGGACATTCAGCTCTTCTGGAGCACTTGGCTGTCAGCATTTGTT	3324
QY	993	LeuArgCysPheLeuGluAlaAspProTyrIleAspIleAspGlnAsnValLeuHisArg	1012
Db	3325	TTAAGAGTCTTTTCGGAAGCTGATTAATATATATATGATACAGATGGTGTACACAGA	3384
QY	1013	ThrTyrThrTrpLeuLysGlyHisGlyLysSerAsnGlyGluPheTrpAspProGlyArg	1032
Db	3385	ACATATATCTGGCTTAATGACATTAAGAAATTCATATGTGATTTTGGAGCCAGGAAGA	3444
QY	1033	ValIleHisSerGluLeuGlnGlyIleAsnLysSerProValThrLeuThrAlaTyrIle	1052
Db	3445	GTGATTCACAGTGAACTTCAAGGGTGACACCAAAAGCCACAGTAAACATTTACGGCTTAATTT	3504
QY	1053	ValThrSerLeuLeuGlyTyrArgLysTyrGlnProAsnIleAspValGlnIleSerIle	1072
Db	3505	GTGACTTCTGTGCTGGGATACAAAAAGTATTCAGCTTAATTCGATGTGACAAAGACTCATATC	3564

QY 1073 HsPheLeuGluSerGluPheSerArgGlyIleSerAspAsnTyrThrLeuAlaLeuIle 1092  
 |||||  
 Db 3565 AAGTTTTCGAATTAATTCACAGAGAAATTCGACAAATTAATACCTTAAGCAATTATA 3624  
 |||||  
 QY 1093 ThrTyrAlaLeuSerSerValGlySerProValAlaGlyAlaLeuMetLeuThr 1112  
 |||||  
 Db 3625 TCTTAATGCGCTGTCACAGTCGGAGCCCTTAAGACAGAGAGGCTTGAACCTTGATG 3684  
 |||||  
 QY 1113 TTPArgAlaGluGlnGluGlyGlyMetGlnPheTyrValSerSerGluSerIleuSer 1132  
 |||||  
 Db 3685 CAGCATCAGAAAGAGAGAGACACACAGTTCTGTGTGTCATCAGAGACTGGCTCTCT 3744  
 |||||  
 QY 1133 AspSerTyrGlnProArgSerSerLeuAspIleGlyValAlaAlaTyrAlaLeuSerHis 1152  
 |||||  
 Db 3745 GGGTCTCTGGCAGCGCTCCGTCGTCGACATCGAAATCGCTGCTTATCAGCTGGCGCAC 3804  
 |||||  
 QY 1153 PheLeuGlnPheGlnThrSerGluGlyIleProIleMetArgTyrLeuSerArgGlnArg 1172  
 |||||  
 Db 3805 ACTCTGCAC--CATGTGTCTGAGGAAATCCGGTATAGAGTGCTCATCCAGCAAGA 3861  
 |||||  
 QY 1173 AasnSerLeuGlyGlyPheAlaSerThrGlnAspThrThrValAlaLeuValAlaLeuSer 1192  
 |||||  
 Db 3862 AACAGCTCGAGAGTTTGTGATCCACTCAGAGACTGTTGTGGCCTTGAAAGCATTAATCT 3921  
 |||||  
 QY 1193 GluPheAlaAlaLeuMetAsnThrGluArgThrAsnIleGlyValThrValThrGlyPro 1212  
 |||||  
 Db 3922 GAATTTTCAGCCCTTGTCACAAAGAAATACAGATATCACTCACTGACCGTCACAGGGCC 3981  
 |||||  
 QY 1213 SerSerPro----- 1215  
 |||||  
 Db 3982 GGCATCCCGAATCCATACACTCCGAATGACTCCAGAACCTCTTCTCTTCCACAG 4041  
 |||||  
 QY 1216 SerProLeuAlaValAlaGlnProMetAlaValAsnIleSerAlaAsnGlyPheGlyPhe 1235  
 |||||  
 Db 4042 GAAAGCTTCATGCGCTAGATCCCATCCAGCCGTTAAATGTTTGGCGATGCTCGGATTT 4101  
 |||||  
 QY 1236 AlaIleCysGlnLeuAsnValValTyrAsnValIleValAsnSerGlySerSerArgArgArg 1255  
 |||||  
 Db 4102 GCTATTTGCCAGCTTAATGTTGCTATTAAGTGAAGGTTTCAAGGTTCTTCTTAAAGACG 4161  
 |||||  
 QY 1256 ArgSerIleGlnAsnGlnGluAlaPheAspLeuAspValAlaValIleGluAsnIleAsp 1275  
 |||||  
 Db 4162 AATATCACCGAAATTCAGAAAGTTTGTGATTAGAGTCATGTCG---AATATGAGGAC 4218  
 |||||  
 QY 1276 AspLeuAsnHisValAspLeuAsnValCysThrSerPheSerGlyProGlyArgSerGly 1295  
 |||||  
 Db 4219 GACATTAGCACCTGAAATCTGAATGTGTGACAGAGTCACTTGGGTTTCAGAGAGGACAGC 4278  
 |||||  
 QY 1296 MetAlaLeuMetGluValAsnLeuLeuSerGlyPheMetValProSerGluAlaIleSer 1315  
 |||||  
 Db 4279 ATGGTCCCTTATGGAAGTGAACCTTCTCAGTGGCTTAAAGTCATCTTCAATTCATTCCT 4338  
 |||||  
 QY 1316 LeuSerGluThrValIleValIleValGluTyrAspHisGlyValLeuAsnLeuTyrLeuAsp 1335  
 |||||  
 Db 4339 CTGAGTGAACCCCTGAAAGAAAGTGAATGTATGAAGGGAACCTCAACCTTATTTAGAT 4398  
 |||||  
 QY 1336 SerValAsnGluThrGlnPheCysValAsnIleProAlaValArgAsnPheIleValSer 1355  
 |||||  
 Db 4399 TCTGTAATGAATCCAGTTTGTGTGTAATATCCACGTGAGACATCAAGAAAGTTTA 4458  
 |||||  
 QY 1356 AsnThrGlnAspAlaSerValSerIleValAspTyrTyrGluProArgArgGlnAlaVal 1375  
 |||||  
 Db 4459 AATATTCAGATGGTTCATATCTGTATGACTACTATGAACCAAGAGACAGAGAGAG 4518  
 |||||  
 QY 1376 ArgSerTyrAsnSerGluValIleValLeuSerSerCysAspLeuCysSerAspValGlnGly 1395  
 |||||  
 Db 4519 CCAAGGTACAAACACCAAGGTGAAGTGTCTCATGTGTACCTCAGTCCAGACAC--AAC 4575  
 |||||  
 QY 1396 CysArgProCysGluAspGlyAlaSerGlySerHisHisSerSerValIlePheIle 1415  
 |||||  
 Db 4576 TGCATATCACACACGAGGAGGACGAGACTCCCTTGAAGCTTTCGAGCCTTCTGTC 4635  
 |||||  
 QY 1416 PheCysPheIleLeuLeuTyrPheMetGlu 1425

Db 4636 TTTTGCTCCGCTCTCTGTGATCTTGCA 4665  
 |||||  
 RESULT 11  
 BC052443  
 LOCUS BC052443 5644 bp mRNA linear ROD 19-MAY-2003  
 DEFINITION Mus musculus cDNA clone MGC:63463 IMAGE:6406841, complete cde.  
 ACCESSION BC052443  
 VERSION BC052443.1 GI:30851465  
 KEYWORDS MGC.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 5644)  
 REFERENCE  
 AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,  
 Klausner,R.D., Collins,F.S., Wagner,L.H., Sherman,C.M., Schuler,G.D.,  
 Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,  
 Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,  
 Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,  
 Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,  
 Scheetz,T.E., Brownstein,M.J., Uedini,T.B., Toshiyuki,S.,  
 Carninci,P., Prange,C., Raha,S., Loquellano,N.A., Peters,G.J.,  
 Abramson,R.D., Miallady,S.J., Bosak,S.A., McEwan,P.J.,  
 McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,  
 Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hultk,S.W.,  
 Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,  
 Fahy,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,  
 Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shevchenko,Y.,  
 Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,  
 Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,  
 Butlerfield,Y.S., Krzywinski,M.I., Skalka,U., Smalke,D.E.,  
 Scherch,A., Schin,J.E., Jones,S.J. and Matre,M.A.  
 Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences  
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
 2 (bases 1 to 5644)  
 12477932  
 JOURNAL MEDLINE  
 PUBMED 22388257  
 REFERENCE  
 AUTHORS Strausberg,R.  
 TITLE Direct Submission  
 JOURNAL Submitted (15-MAY-2003) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 CONTACT Contact: MGC help desk  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: Dr. Jim Lin, University of Iowa  
 cDNA Library Preparation: M. Bento Soares, University of Iowa  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)  
 DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr.  
 Thomas L. Casavant.  
 Web site: <http://genome.uiowa.edu>  
 Contact: [bento-soares@uiowa.edu](mailto:bento-soares@uiowa.edu); [tom-casavant@uiowa.edu](mailto:tom-casavant@uiowa.edu)  
 Bonaldo,M.F., Akabogu,I., Bair,T., Bair,J., Crouch,K., Davis,A.,  
 Fisher,K., Keppel,C., Kucaba,T., Lebeck,M., Melo,A., Schaefer,K.,  
 Scheetz,T., Smith,C., Shirk,E., Tack,D., Trout,K., Walters,J.,  
 Casavant,T., Soares,M.B.  
 Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LINL at: <http://image.lnln.gov>  
 Series: Place: Row: Column: 0.  
 Location/Qualifiers  
 1. 5644  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6"  
 /db\_xref="taxon:10090"  
 /clone="MGC:63463 IMAGE:6406841"  
 /tissue\_type="Brain, enriched mouse brain 12.5dp"  
 /clone\_lib="NIH BMAP\_F00"  
 /lab\_host="DHI0B"

## FEATURES

## source



CDS

```

/Note="Vector: pYX-ASC"
105..4433
/codon_start=1
/product="unknown (protein for MGC:63463)"
/protein_id="AAH52443.1"
/db_xref="GI:30851466"
/translation="MRSRRILISAHLILCLCAVALAAGSRFLVTAPGIIIRPGANVTIG
VDLENSPQVLVAQVFKIASNKRSLILAEVFNHGRKTLVLPALPLPSADKLYE
LHNGOSENEIVFSNTRLTPTESKSIIVLTQDPAFKPQEVFRVLTLCDLKPSR
TSVDIPIKDKPSNVIQOMFSQKGLVAGSNIPDPSNPIQVNVDOOYQSPF
OVLEYVLPRFEVTVQPLCYSKSKOLNGSVIAKYTGKVGKSLSTPLPSFMGKK
KITKSPREINGFANFSDNYEMKVMKLKPLVSESSYENDPSFPGRAEITATVTE
SLTGISMASTNVFFKQHDYIIEIFDTVLKSLNLTAVKSRSDGNOLTPREIN
DLVTVAQRKNHNEFSQDQMDYIQVNTYIIPONGIKIEFPVMSISGELQKAYFL
DGTSSVTHSMFTSPSKTYIQLKTRDEYIVAGSPDLMVGNQKOLSYMVISKGL
VAGKOSRTFSLTPEASMAKACIIVAYIAEDEIINDILKIPIVOLVFNKYLFW
KPTKPSDKSLRISATOSDSIVIGVAVDSVILMENSITMETWHELEKNTXYE
LGMFMSFAVQECGLMVLDTATIRSIDVYDTEYSRPAENANLVDEDDASS
VNVVRKRPETWILWDATMGSKITSEFEVTPDSITSVASAFVISEDGLGLTV
PAVLQAFQPFPLFLNLPYSVIRGEFALVSVLYNLTQILVILIBESDPILMTS
NDNGTIYRKTVQVPRDNGVTLVPEIPKTHGEIPITVTAASPTASDAVQTIIVPE
GIEKSYKSVLLDLTDSNVEKQOOSMSEFPDPTVIGSEVOITAGIDIGSINGIS
SLIRMPYCGEONMIVFAPNVIYLDVLTOKOLTVNKEKALSYMROGORELLXORE
DGSFARQDIDSSGSTVLSAFVLRCEFLADYIIDDDVLRHTYMLNAHKNGEFLM
EPGRVHSELDQGTSPVTLTAIVTVSVLKYQPIIDVQDSIKFLBEFSSPDIEN
YTTAIIISYALSTVGSFPAEBALNLMORSEKEDTQFMLSISGMOFPRVDIEI
AAVALAHATHVSEGIPIVWRMLIQONSLSGFVSTQDITVVALKALSEFALVKE
DIQLTVTGPGIPIRSIHRIDSONFLHOBELHALDITVNSAHSGFALCOLNDY
NVGSGSKRRSRSTENOEVFDIYVNNMEDISHLNNTCTSHLSGSRGMVMEVNL
LSGFSASDSIPLSETIKAKYEDNGKNTLYLDSVNSQFCVNIPTVADIVSVNTRGS
VSVMDYEPFRQAVRSNTYVKLSSCYLSPDINCKSHTDATDSLRSSSLVLFCSVL
LYPVQHY"

```

BASE COUNT 1556 a 1235 c 1251 g 1602 t

ORIGIN

Alignment Scores:

Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
0	5360.00	83.72%	71.59%	72.95%	5644	1038	176	202	34	8

US-10-020-095-4 (1-1428) x BC052443 (1-5644)

QY 1 MetGlnGlyProProlLeuThrAlaAlaHisLeuLeuCysValCysThrAlaAlaLeu 20

Db 105 ATGCGCAGTGGAGGCTCTGAGCGGCCCACTTCTGCTGTGCGCGCTGCGCTG 164

QY 21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGlyGly 40

Db 165 GCC---GCCCAAGGCTCCGCTTCTGTGTGACAGCCCAAGGATCATCAACCCGAGCA 221

QY 41 AsnValThrIleGlyValGluLeuLeuGluHisCysProSerGlnValThrValValAla 60

Db 222 AATGTGACTTGTGGGTGAGATCTCTCGAAAAATPAGCCCCACAGGCTCTGTGAAGGCT 281

QY 61 GluLeuLeuYsrThrAlaSerAsnLeuThrValSerValLeuGluGluGluValPhe 80

Db 282 CAGGTGTCAAGATAGCTTCCAAACAAATCGAATCCATCTCGAAGACGAGAGGCTTTT 341

QY 81 GluIlyGlySerPheYsrThrLeuThrLeuProSerLeuProlLeuAsnSerAlaAspGlu 100

Db 342 CACAGAGGCAATTCAAGACTCTCGTCTCCGCGCACTACCTTGACGACATGCAAGTAAG 401

QY 101 IleTyrGluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThr 120

Db 402 ATTATGAGCTACACATAACGCAATCAGAGATAGATCTATCTTCCACAGACGACA 461

QY 121 ArgLeuSerPheGluThrIlySerArgIleSerValPheIleGlnThrAspIlyAlaLeuTyr 140

Db 462 CGTTAACATTATGAGCAAGATATATCTGCTCTCACTTCAACAGACGAGCAATTCTAC 521

QY 141 LysProlGlnGluValIlySerPheArgIleValThrLeuPheSerAspPheYsrProTyr 160

Db 522 AACCCAGAGGAGAGGAAATTTCCGGGAGCTCCACACTGCTCAGATCTCAAGCCTTAT 581

QY 161 LysThrSerLeuAsnIleLeuIleIlyAspProlYsrSerAsnLeuIleGlnGlnTyrPhe 180

Db 582 AGACACATCCGTGACATTTTCATTAGATCCCAATCTCAATATGATCCCAACGAGTGT 641

QY 181 SerGlnIleSerPheGluValIleSerTyrPheGlnLeuSerSerHisProlIle 200

Db 642 TCTCAGAAAGATATCTTGAGAGCTGTTCCAAACCTTTTCAACTATCTTAACTCAATA 701

QY 201 LeuGlyAspTyrPheIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGln 220

Db 702 TTGGTGACTGGGCCATTCAAGTTCAAGTAAGATCAGCAATATTAACGATCATTTGAG 761

QY 221 ValSerGluTyrValLeuProlYsrPheGluValThrLeuGlnThrProlLeuTyrCysSer 240

Db 762 GTCTAGAAATATGTGTATCAAGATTCAGAACGTCACCGTCAGACCCCGCTGATTTGCTC 821

QY 241 MetAsnSerIlyHisLeuAsnGlyThrIleThrAlaIlySerTyrThrTyrGlyLysProlVal 260

Db 822 TTGAAGTCTTAACAGTTAAATGCTCCGTCATAGCAAAATATATACGAGAACGAGT 881

QY 261 LysGlyAspValThrLeuThrPheLeuProlLeuSerPheTyrPheIlyLysIlyAsnIle 280

Db 882 AAGGAAAGCTGTGACATTCATTATTAACCTTTGTTGGGAAAGAAAGAAATATATC 941

QY 281 ThrIlyThrPheIlyIleAsnGlySerAlaAsnPheSerPheAsnAspGluGluMetLys 300

Db 942 ACAAAGATTTTGAGATTAATGATTTGCAAACTCTCTTGTGATATATATGATGAGAA 1001

QY 301 AsnValMet-----AspSerSerAsnGlyLeuSerGluTyrLeuAsp 314

Db 1002 AAGTATATGAATTTGAAACCCGTAACGATGCTCCGAAGGAGATTATGAAATGTGAC 1061

QY 315 LeuSerSerProGlyProValGluIleLeuThrThrValThrGluSerValThrGlyIle 334

Db 1062 CGGTCTTTTCCGGGCGGAGAGATTAAGTACTGACGGAATCAGCTCAGAGTATC 1121

QY 335 SerArgAsnValSerThrAsnValPhePheIlyGlnHisAspTyrIleIleGluPhePhe 354

Db 1122 TCAGATGACCAAGCAACCAATGTGTTTTCAGCAACATGATTAACATGATGATTTT 1181

QY 355 AspTyrThrThrValLeuYsrProSerLeuAsnPheThrAlaThrValLysValThrArg 374

Db 1182 GATTATACAGACTGTCTTGAACCATCTCTCAACTTACAGCCACTGTGAAGTCAAGCCG 1241

QY 375 AlaAspGlyAsnGlnLeuThrLeuGluGluArgAsnAsnValIleThrValThr 394

Db 1242 TCTGATGCAATCAACTACTCTCGAAGAAATAGAAATGATTTGTGTCAGATGATGACC 1301

QY 395 GlnArg-----AsnTyrThrGluTyrTyrSerGlySerAsnSerGlyAsnGlnLysMet 412

Db 1302 CAGCGGAAAAATACACACCTGAG-----AGTCAGACAGATCAGGAAAG 1346

QY 413 GluAlaValGlnIlyIleAsnTyrThrValProGlnSerGlyThrPheIlyIleGluPhe 432

Db 1347 GATTATATCAAGCCGTAATATATATATATATATATATATATATATATATATATAT 1406

QY 433 ProlLeuGluAspSerSerGluLeuGlnLeuLysAlaTyrPheLeuGlySerLysSer 452

Db 1407 CCAAGTATTCATTTCCGGGAGGCTACACAGAGGCTTATTTCTTGAATGATGACAAAT 1466

QY 453 SerMetAlaValHisSerLeuPheYsrProSerLysThrTyrIleGlnLeuLysThr 472

Db 1467 TCTGTGACAGTCCACAGTATGTTCACCTCTCTAGTAAGACCTACATCCAGCTTAAAC 1526

QY 473 ArgAspGluAsnIleIlySerValGlySerProlGluLeuValValSerGlyAsnLysArg 492

Db 1527 AAGAGATGAATATATTAAGGTGGGCTACCTTTTGAATTTGATGTGAGCAACGACAA 1586

QY 493 LeuLysGluLeuSerTyrMetValValSerArgGlyGlnLeuValAlaValGlyLysGln 512



```

Db      1587 TTCAAGGACTTAAGCATATGTAATGTAATCAAAAGGACAGTGGCTGCAGAGCAAAACA 1646
Qy      513 AsnSerThrMetPheSerLeuThrProGluAsnSerTrpThrProValAlaCysValIle 532
Db      1647 AGTTCAGAAAGACTTCTCTTTAACAACAGAGGCTTCTGGGCTCCAAAGGCTGTATAAATC 1706
Qy      533 ValTyrTrilleGluAspArgIleGluIleIleSerAspValLeuValIleProValGln 552
Db      1707 GCATATTATATTCAGAGGATGGGAAATTATAATGATATCTTAATAAATCCCGCTTCAG 1766
Qy      553 LeuValPheValAsnValIleValLeuTyrTrpSerLeuValIleValProSerGlu 572
Db      1767 CTGTGTTTGAATAATAGTAAAGCTGTTTGAAGTAAACCTACAGTCAAGCATCTGAT 1826
Qy      573 LysValSerLeuArgIleSerValThrGlnProAspSerIleValGlyIleValAlaVal 592
Db      1827 AAGGTCTCTCTCGAGATCTCTGCGACACAGTCTGACTCCCTGGTGGGAGTGTAGCTGTT 1886
Qy      593 AspLysSerValAsnLeuMetAsnAlaSerAsnAspIleThrMetGluAsnValAlaHis 612
Db      1887 GACAAAGAGTAACTTGAATGAGAAATCTCCACAGATTTACATGAGAAACCATGCTCAT 1946
Qy      613 GluLeuGluLeuTyrAsnThrGlyTyrTyrLeuGlyMetPheMetAsnSerPheAlaVal 632
Db      1947 GAAGTGAAGCTTATTAACACAGAAATATTAATCTAGGACATGTCATGAATCTTCGCTGTA 2006
Qy      633 PheGlnGluCysGlyLeuThrValLeuThrAspAlaAsnLeuThrIleAspValIleAsp 652
Db      2007 TTTCAGGAATGGTGTCTGGGTGGTGAAGATGCAACCTTATTAAGATTAAGCATTTGAT 2066
Qy      653 GlyValTyrAspAsnAlaGluTyrAlaGluArgPheMetGluGluAsnGluGlyHisIle 672
Db      2067 GAAGTACAGATACGTAAGAGTATCTGAACGCTTTCAGAGAGAAATAGAGCAAACTCG 2126
Qy      673 ValAspIleHisAspPheSerLeuGlySerSerProHisValArgLysHisPheProGlu 692
Db      2127 GTTGATTTTGAAGATCTTCTTCAGTTAAACAATGTACATGTACGAGAAGATTTTCCAGA 2186
Qy      693 ThrTrpIleLeuLeuAspThrAsnMetGlyTyrArgIleTyrGlnGluPheGluValIle 712
Db      2187 ACATGATTTGGTGTACACGCTTACATGAGATCTTAACATTTATGCAAAATTTGAAAGTTAC 2246
Qy      713 ValProAspSerIleThrSerTrpValAlaThrGlyPheValIleSerGluAspLeuGly 732
Db      2247 GTRACGACTCCATCACTTCTTGGGTGGCTTGGCTTTGTCACTCAAGAAAGCTCGGAT 2306
Qy      733 LeuGlyLeuThrThrThrProValGluLeuGlnAlaPheGlnProPhePheIlePheLeu 752
Db      2307 TTGGGTTTAACAACCGTTCAGACAGAGCTGCMAAGCCTTCCAAACGTTTTCCTTTCCG 2366
Qy      753 AsnLeuProTyrSerValIleArgIleGluGluPheAlaLeuGluIleThrIlePheAsn 772
Db      2367 AATCTTCCGTACTCTGTATCAAGGTGAAGAGTTTGGAACTTCAATCCATCCGCAAT 2426
Qy      773 TyrLeuLysAspAlaThrGluValLysValIleIleGluLysSerAspLysPheAspIle 792
Db      2427 TATTTAAAGATTAACAATTAAAGTTGTGATCACTATGAGAAAGTAAACAGCTTGTATAT 2486
Qy      793 LeuMetThrSerSerGluIleAsnAlaThrGlyHisGlnGlnThrLeuLeuValProSer 812
Db      2487 TTATATGACTTCAAAATGACACCAATGACACCATATACCGGAAGCTTCAAGTTCCACAG 2546
Qy      813 GluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGlyGluIleProIle 832
Db      2547 GATTAAGGGGTCACTGTTTTCATCAACGCAACCACTGGGAGAGATTTCCCATC 2606
Qy      833 ThrValThrAlaLeuSerProThrAlaSerAspAlaValThrGlnMetIleLeuValLys 852
Db      2607 ACGATACGGCTCCCACTCCACTGCTGTAGTGTCAACCAAGCAATTTGATGAAG 2666
Qy      853 AlaGluGlyIleGluLysSerTyrSerGlnSerIleLeuLeuAspLeuThrAspAsnArg 872

```

```

Db      2667 CCTGAAGGAGTAAAGAGCTGATTCGAATCTGTCTTATTTGGATTCAGATAGCAAC 2726
Qy      873 LeuGlnSerThrLeuLeuThrLeuSerPheSerPheProProAsnThrValThrGlySer 892
Db      2727 GTAGAAAGTAAACACCAATCTATGAGATTTCTTCTTCTCTGATGAGGTCAATTTGGAGT 2786
Qy      893 GluArgValGlnIleThrAlaIleGlyAspValLeuGlyProSerIleAsnGlyLeuAla 912
Db      2787 GAAAGGTTCAATCAACAGCAATTTGAGATATCTTGGTCTCTCATCAATGGCTTATCT 2846
Qy      913 SerLeuIleArgMetProTyrGlyCysGlyGluGluAsnMetIleAsnPheAlaProAsn 932
Db      2847 TCACGTAGTCCGGATCCCTTATGATGATGTGTGAACGAACATGATATATTTTCTCCAAT 2906
Qy      933 IleTyrIleLeuAspTyrLeuThrIleValLysValGlnLeuThrAspAsnLeuLysGluLys 952
Db      2907 ATTTCATTTTGGATTTATCTGACTAAACAGAAACGCTGACGTTAAATTTAAAGAAAG 2966
Qy      953 AlaLeuSerPheMetArgGlnGlyTyrGlnArgIleLeuLeuTyrGlnArgGluAspGly 972
Db      2967 GCCCTTCAATATAGGCAAGTTACCAAGAGAGCTTCTCTATCAAGAGGAGATGGC 3026
Qy      973 SerPheSerAlaPheGlyAsnTyrAspProSerGlySerThrTrpLeuSerAlaPheVal 992
Db      3027 TCCTTCAGTGTCTTTGGGAGCAATGACTTCTGCGAACACTGGCTGTCAACATTTGTT 3086
Qy      993 LeuArgCysPheLeuGluAlaAspProTyrIleAspIleAspGlnAsnValLeuHisArg 1012
Db      3087 TTAAGATCTTCTTGAAGCTGATTAATTAATGATATTAATGATGATGTGTACACAGA 3146
Qy      1013 ThrTyrThrTrpLeuLysGlyHisIleGlnLysSerAsnGlyGluPheTrpAspProGlyArg 1032
Db      3147 ACATATCTCTGGCTTAATGACATTAAGAAATTCATATGTTGATTTGGAGCCAGGAAGA 3206
Qy      1033 ValIleHisSerGluLeuGlnGlyValLysAsnLysSerProValThrLeuThrAlaTyrIle 1052
Db      3207 GTGATTTCAAGGAATCTCAAGGTGGCACCAAAAGCCAGTAACCTTACCGGCTATAT 3266
Qy      1053 ValThrSerLeuLeuGlyTyrArgLysTyrGlnProAsnIleAspValGlnGluSerIle 1072
Db      3267 GTGACTTCTGCTCGGTGAACAAAAGATACAGCTTAATATGATGATGATGATCAACATCAATC 3326
Qy      1073 HisPheLeuGluSerGluPheSerArgGlyIleSerAspAsnTyrThrLeuAlaLeuIle 1092
Db      3327 AAGTTTGGATTTGAATTCAGACAGAGAAATTCGCAATTAATCACTTACCAATTATTA 3386
Qy      3387 ThrTyrAlaLeuSerSerValGlySerProLysAlaLysGluAlaLeuAsnMetLeuThr 1112
Db      3387 TCCTATGCCGTGTCCACAGTGGAGCCCTTAAGACAGAGAGGCTTGAACCTGCTGATG 3446
Qy      1113 TrpAlaGluGlnGlnGlyGlyMetGlnPheTrpValSerSerGluSerLysLeuSer 1132
Db      3447 CAGCATCAAGAAAGAGAGAGACACACAGATTCGTGTGTATCATCAGACCTGCGCTCTCT 3506
Qy      1133 AspSerTrpGlnProArgSerLeuAspIleGluValAlaAlaTyrAlaLeuLeuSerHis 1152
Db      3507 GGGCTTCGACAGCCGCTCCGTGACATCGAAATCGGTGCTTATGACCTGTGGCGAC 3566
Qy      1153 PheLeuGlnPheGlnThrSerGluGlyIleProIleMetArgTrpLeuSerArgGlnArg 1172
Db      3567 ACTCTGAC---CATGTGTGAGAGGAAATCCGGTTATGAGGTGCTCATCAGCAAAAGA 3623
Qy      1173 AsnSerLeuGlyGlyPheAlaSerThrGlnAspThrThrValAlaLeuLysAlaLeuSer 1192
Db      3624 AACACCTCGGAGGTTTGTATCACTAGAGATCTGTGTGGCCCTGAAGGCATTAATCT 3683
Qy      1193 GluPheAlaAlaLeuMetAsnThrGluArgThrHisIleGlnValIleThrValThrGlyPro 1212
Db      3684 GAATTTTCAAGCTTGTGTGACAAAGAAATTAACATATTCACATGACCGTCAACAGGCC 3743
Qy      1213 SerSerPro----- 1215
Db      3744 GGCATCCCAAGATCACTACACTTCGAAATGATGATCCAGAAACCTTTCTCTTCCACAG 3803

```

```

Qy 1216 SerProLeuAlaValAlaGlnProMetAlaValAsnIleSerAlaAsnGlyPheGlyPhe 1235
Db 3804 GAAGAGCTTCATGGCGCTGATCCCATTCACCGTTATGTTTCTGGCAGATGGCTCGGAAATT 3863
Qy 1236 AlaIleCysGlnLeuAsnValAlaIleAsnValIleValAlaSerGlySerSerArgArgArg 1255
Db 3864 GCTATTTGCGCAGCTTAATGTTGACTATTAACGTGAAGGTTTCTTCTTAAAGACGG 3923
Qy 1256 ArgSerIleGlnAsnGlnGluAlaPheAspLeuAspValAlaValIleGluAsnIleAsp 1275
Db 3924 AGATCTACCGAAATCAAGAAAGTTTTCATTTAGACGTCATTGG--AATTAATGAGGAC 3980
Qy 1276 AspleuAsnIleValAspLeuAsnValCysThrSerPheSerGlyProGlyArgSerGly 1295
Db 3981 GACATTAGTCAACCTGAATCTGAATGTGCAACAGTCACTGGGTTCAGAGAGACAGGC 4040
Qy 1296 MetAlaLeuMetGluValAsnLeuLeuSerGlyPheMetValProSerGluAlaIleSer 1315
Db 4041 ATGGCTCTTATGGAAGTGAACCTTCTCAGTGGCTTACTGATCTTCAGATTCAATTCT 4100
Qy 1316 LeuSerGluThrValIleValIleValIleValIleValIleValIleValIleValIleVal 1335
Db 4101 CTGAGTGAACCTCGAAGAAAGTGAATATGATTAACGGAATCTCAACCTTTATTTAGAT 4160
Qy 1336 SerValIleAsnGluThrIlePheCysValAsnIleProAlaValAlaArgAsnPheIleValSer 1355
Db 4161 TCTGTAATGAATCCAGTCTTGTGTTATATTTCCCACTGAGAGACTCAAAAGTTTCA 4220
Qy 1356 AsnThrGlnAspAlaSerValSerIleValAspIleValIleValIleValIleValIleVal 1375
Db 4221 AATTTTCAGATGGTTCAGTATCTGTAATGACTACTATGAACCAAGAGACAGGACAGTG 4280
Qy 1376 ArgSerIleAsnSerGluValIleValIleValIleValIleValIleValIleValIleVal 1395
Db 4281 CGAAGCTACCAACCCAGGTGAAGTGTCTCAATGATTAACCTCAAGTCCAGAAC--AAC 4337
Qy 1396 CysArgProCysGluAspGlyAlaSerGlySerIleHisIleSerValIlePheIle 1415
Db 4338 TGCAGATCACACAGGACGAGACGAGCAGCCTCCCTGAGACTTCTTGAGGCTTCTGTC 4397
Qy 1416 PheCysPheIleValLeuIleValIleValIleValIleValIleValIleValIleVal 1425
Db 4398 TTTTGCTCCGCTTCTGTTACTTGTGCAA 4427

```

```

RESULT 12
HSM805581      2938 bp      mRNA      linear      PRI 12-JUN-2002
LOCUS          Homo sapiens mRNA; cDNA DKFZp762L1111 (from clone DKFZp762L1111).
DEFINITION     AL834478
ACCESSION      AL834478.1 GI:21740252
VERSION        1
KEYWORDS       Homo sapiens (human)
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2938)
REFERENCE      Blum, H., Bauserachs, S., Mewes, H.W., Weil, B. and Wiemann, S.
AUTHORS       Direct Submission
TITLE          Submitted (09-JUL-2002) 1, D-85764 Neuberberg, GERMANY
JOURNAL        Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
COMMENT        Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
                sequenced by LMU (Ludwig Maximilians University,
                Munich/Germany) within the cDNA sequencing consortium of the German
                Genome Project.
                This clone (DKFZp762L1111) is available at the RZPD in Berlin.
                Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
                Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
                information about the clone and the sequencing project is available
                at http://mips.gsf.de/proj/cDNA/.
                Location/Qualifiers
                1..2938
                /organism="Homo sapiens"

```

```

FEATURES
Source
1..2938
/organism="Homo sapiens"

```

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/map="15"
/clone="DKFZp762L1111"
/tissue_type="melanoma (Mewo cell line)"
/clone_lib="762 (synonym: hmel12). Vector pSport1; host
DH10B; sites NotI + SalI"
/dev_stage="adult"
1..2938
/gene="DKFZp762L1111"
<1..2565
/gene="DKFZp762L1111"
/note="similarity to alpha-2-macroglobulin (Limulus sp.)."
/codon_start=1
/product="hypothetical protein"
/protein_id="CAD39137.1"
/db_xref="GI:21740253"
/translations="VRSVNLNMAASDITMENVVEHELELYNTGYLGMPNRAVQVE
CGLWVLDAVLTDYIDGVYDNEAVERFMEVEEGHVIDHDSLSSPVRKHPET
MILDTNMGSRIVQREVEVTPDSITSMVATGVISDGLTTEVELOAQPFI
LMLPSVIGEEFALIEITIFNYLKDAEYVILIEKDFDILMTSEINATGQOFL
VPEEDCATVLFPIRPHLGEIPITVVALSPASDAITOMILVKAEGIEKSYOSILD
LTNRIOSLTXTLSFSPRPNTVSGSRVOTITADVLGSLINGSLALIMPVCGGN
MIRPANIYITLITLTKKQITDNLKALSPKQYQRELYREGSRSAPNTYPS
GSTWLSAFVLRCLFADPYIIDQNLHRTYTMKQHQSNGSEFMDPGRVHSELQGS
NKSPVLTAVITVTSIGYRKYOPNIDVQSIHLESEFSGISDNYTLALITYALSQ
GSPKAKELNMLTAEQGMQFVSSKLSDSQPSRLDIEVAVALSHFLQFQ
TSRGIPIRMRLSGSRNLSGSPASTODTVALKALSEFALMSEFALMSEFALMSEF
SPYKELIDHNRILLQTAELAVQPTAVNISANGRPALCOLNVVNVKVASGSR
SINQRAFDIDAVKENDKODLVNVDVCTSRSGSPSGMLMEVNLISGFWPSAI
SLSEYKVEYDHGKMLNLYDSVNETQFCVNI PAVNFPVSNQDASVIVDYEEPR
QAVRSYNSREYKLSCDLCSDVQCRDEBGASGSHHSVIFPCFLXYFMELWL"
2843..2848
/gene="DKFZp762L1111"
2859
polyA_site
2859
/gene="DKFZp762L1111"
BASE COUNT      887 a      529 c      620 g      902 t
ORIGIN
Alignment Scores:
Pred. No.:      5.54e-274      Length:      2938
Score:          4305.50      Matches:      834
Percent Similarity: 97.78%      Conservative: 1
Best Local Similarity: 97.66%      Mismatches: 2
Query Match:    58.59%      Indels:      17
DB:              9      Gaps:      1
US-10-020-095-4 (1-1428) x HSM805581 (1-2938)
Qy 592 ValAspIleSerValAsnLeuMetAlaValAsnAspIleThrMetGluAsnValAl 611
Db 1 GTTGACAAAGTGTGAATCTGATGAATGCCCTCAATGATTTCAATGCAAAATGTGCTC 60
Qy 612 HisGluLeuGluLeuIleValIleValIleValIleValIleValIleValIleValIleVal 631
Db 61 CATGAGTTGCACTTATTAACACAGATATATTATTTGCGCATGTTCTTATGAAATTTTGGCA 120
Qy 632 ValPheGluGluGlyGlyLeuThrValIleValIleValIleValIleValIleValIleVal 651
Db 121 GTCTTCAGAAAGTGAAGTCTGGATTTGACAGAACTCAAGAAAGATTATTT 180
Qy 652 AspGlyValIleValAspAsnIleValIleValIleValIleValIleValIleValIleVal 671
Db 181 GATGGCTTTATGACAAATGAGAAATGCTGAGAGGTTTATGAGGAAATGAGAGCAT 240
Qy 672 IleValAspIleHisAspPheSerIleGlySerSerProHisValArgIleHisPhePro 691
Db 241 ATTTGATATTTGATGATCTTTCTTTGGTAGAGTCCACATGTCGAAAGCATTTTCCA 300
Qy 692 GluThrThrIleThrLeuAspThrAsnMetGlyIleArgIleIleValIleValIleValIleVal 711
Db 301 GAGACTTGATTTGGCTAGACACCAACATGGGTTCCAGAGATTATCAAGAAATTGAGATA 360

```

QY 712 ThrValProAspSerIleThrSerTrpValAlaThrGlyPheValIleSerGluAspLeu 731  
 Db 361 AcGTACCTGATTCATCTCTGCGGCTACAGGTGTTGTGATCTCGAGGACCTG 420  
 QY 732 GlyLeuGlyLeuThrThrThrProValGluLeuGlnAlaPheGlnProPheIlePhe 751  
 Db 421 GGCCTTGACTAACCACTACTCCAGTGGAGCTCCAAAGCTTCCAAACATTTTTCATT 480  
 QY 752 LeuAsnLeuProTyrSerValIleArgGlyGluGluPheAlaLeuGluIleThrIlePhe 771  
 Db 481 TTGAATCTTCCCTACTCTGTATCAGAGGGAGAAATTTGCTTGGAAATTAATATATTC 540  
 QY 772 AsnTyrLeuLeuAspAlaThrGluValLysValIleIleGluLysSerAspLysPheAsp 791  
 Db 541 AATTAATTTGAAAGATGCCACTGAGGTAAAGTAATCATGAGAAAGTGCAAAATTTGAT 600  
 QY 792 IleLeuMetThrSerSerGluIleAsnAlaThrGlyHisGlnGlnThrLeuLeuValPro 811  
 Db 601 ATTCTAATGACTTCAAGTGAATAAATGCAAGGCGACAGCAGCAACCTTCTGATGCC 660  
 QY 812 SerGluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGlyGluIlePro 831  
 Db 661 AGTGAAGATGGGGCAACTGTTCTTTTCCATCAGGCGCAACATCTGGAGAGAAATTCCT 720  
 QY 832 IleThrValThrAlaLeuSerProThrAlaSerAspAlaValThrGlnMetIleLeuVal 851  
 Db 721 ATACAGCTCACAGCTCTTTCACCCAGCTTCTGATGCTATCACCCAGATGATTTAAGTA 780  
 QY 852 LysAlaGluGlyIleGluLysSerTyrSerGlnSerIleLeuLeuAspLeuThrAspAsn 871  
 Db 781 AAGGCTGAAGAAATGAGAAATCATATCATCAATCAATCTTAATTAACCTGATGCAAT 840  
 QY 872 ArgLeuGlnSerThrLeuLysThrLeuSerPheSerPheProProAsnThrValThrGly 891  
 Db 841 AGGCTACAGAGTACCTCGAAAACCTTGAAGTTCTCATTTCTCTTAATCAGAGCTGGC 900  
 QY 892 SerGluArgValGlnIleThrAlaIleGlyAspValLeuGlyProSerIleAsnGlyLeu 911  
 Db 901 AGTGAAGAAGTTCAGATCATCTGCAATGAGAGATGTTCTTGTCCTTCCATCAATGCTTA 960  
 QY 912 AlaSerLeuIleArgMetProTyrGlyGlyGlyGluGlnAsnMetIleAsnPheAlaPro 931  
 Db 961 GGCCTATGATGATGAGATGCTTAATGCTGTGGTGAACAGAAACATGATTAATTTGCTCCA 1020  
 QY 932 AsnIleTyrIleLeuAspTyrLeuThrLysLysLysGluLeuThrAspAsnLeuLysGlu 951  
 Db 1021 AATATTACATTTTGGATTAATCTGACTMAAAAGAAACAACCTGCAACATTAATTTGAAAGA 1080  
 QY 952 LysAlaLeuSerPheMetArgGlnGlyTyrGlnArgGluLeuLeuTyrGlnArgGluAsp 971  
 Db 1081 AAGGCTCTTTCATTTATGAGGCAAGGTACACAGAGAACTTCTCATCAGAGGGAAGAT 1140  
 QY 972 GlySerPheSerAlaPheGlyAsnTyrAspProSerGlySerThrTrpLeuSerAlaPhe 991  
 Db 1141 GGCCTCTTCAAGTCTTTTGGAAATTAAGACCTTCTGAGGAGCACTGATGTCAGCTTTT 1200  
 QY 992 ValLeuArgCysPheLeuGluAlaAspProTyrIleAspIleAspGlnAsnValLeuHis 1011  
 Db 1201 GTTTTAAGATGTTCTCTGAAGCCGATCTTACATGATATGATCAGAAATGTGTTACAC 1260  
 QY 1012 ArgThrTyrThrTrpLeuLysGlyHisGlnLysSerAsnGlyGluPheTrpAspProGly 1031  
 Db 1261 AGAACAATACCTTGGCTTAAAGGACATCAGAAATCCACGCTGAATTTTGGAGATCCAGA 1320  
 QY 1032 ArgValIleHisSerGluLeuGlnGlyGlyAsnLysSerProValIleThrLeuThrAlaTyr 1051  
 Db 1321 AGGTATATCATGATGAGCTTCAAGGTGGCAATMAAGTCCAGATACTTACAGCTTAT 1380  
 QY 1052 IleValThrSerLeuLeuGlyTyrArgLysTyrGlnProAsnIleAspValGlnGluSer 1071  
 Db 1381 ATTGTAACTTCTCTCTCGGAGATTAAGAAAGTATCAGCTTAATATGATGCAAGAGCTCT 1440  
 QY 1072 IleHisPheLeuGluSerGluPheSerArgGlyIleSerAspAsnTyrThrLeuAlaLeu 1091

Db 1441 ATCCATTTTGGAGCTCGAATTCAGTAGAGAAATTTTACAAATTAATTAATCTGACCTT 1500  
 QY 1092 IleThrTyrAlaLeuSerSerValGlySerProLysAlaLysGluAlaLeuAsnMetLeu 1111  
 Db 1501 ATAATCTTATGATTTCTCATGCTGGGAGTCTTAAGCGAAAGAGCTTTGATATATGCTG 1560  
 QY 1112 ThrTrpArgAlaGluGlnGluGlyGlyMetGlnPheTrpValSerSerGluSerLysLeu 1131  
 Db 1561 ACTTGAGAGGAGCAACAAAGAAAGTGGCATGCAATTCGGGTGTATCAGAGTCCAAATCT 1620  
 QY 1132 SerAspSerTrpGlnProArgSerLeuAspIleGluValAlaAlaTyrAlaLeuLeuSer 1151  
 Db 1621 TCTGACTCTCGGACACCGCTCCCTGGATTTGAAATGGACGCTTATGCACTGCTTCA 1680  
 QY 1152 HisPheLeuGlnPheGlnThrSerGluGlyIleProIleMetArgTrpLeuSerArgGln 1171  
 Db 1681 CACTTCTTACAAATTTTCAACTTCTGAGGGAATCCCAATTAATGAGGTGGCTAAGCGCAA 1740  
 QY 1172 ArgAsnSerLeuGlyGlyPheAlaSerThrGlnAspThrThrValAlaLeuLysAlaLeu 1191  
 Db 1741 AGAATTAAGCTTGAGTGGTTTGTCATCTACAGATACCACTGTGGCTTTAAAGCTCTG 1800  
 QY 1192 SerGluPheAlaAlaLeuMetAsnThrGluArgThrAsnIleGlnValIleThrValThrGly 1211  
 Db 1801 TCTGAATTTGGACGCCCTAATGAAATACAGAAAGCAAAATATCCAAGTAGACGTGCGGG 1860  
 QY 1212 ProSerSerProSerPro----- 1217  
 Db 1861 CCTAGCTCACCAAGTCTGTAAAGTTTCTGATGACACACACAAACCGCTTACTCTTCAG 1920  
 QY 1218 -----LeuAlaValAlaGlnProMetAlaValAsnIleSerAlaAsnGlyPheGly 1234  
 Db 1921 ACAGAGACGCTTGTGTGTGACAGCCACGCGAGTAATATTTCCGCAAAATGTTTGA 1980  
 QY 1235 PheAlaIleCysGlnLeuAsnValAlaTyrAsnValLysAlaSerGlySerSerArgArg 1254  
 Db 1981 TTTGCTAATTTGCACTCAATGTTGTATATATGAAAGGCTTGTGGCTCTTCTGAAGA 2040  
 QY 1255 ArgArgSerIleGlnAsnGlnGluAlaPheAspLeuAspValAlaValLysGluAsnLys 1274  
 Db 2041 CGAAATCTATTCAAAAATCAAGAAAGCTTTGATTAATTAATGTTGCTGAAGAAATATAA 2100  
 QY 1275 AspAspLeuAsnHisValAspLeuAsnValCysThrSerPheSerGlyProGlyValArgSer 1294  
 Db 2101 GATGATCTCAATCATGTGATTTGATGTGTACAACTTTTGGGCGCGGATGAGAT 2160  
 QY 1295 GlyMetAlaLeuMetGluValAsnLeuLeuSerGlyPheMetValProSerGluAlaIle 1314  
 Db 2161 GGCATGGGCTTATGGAAGTTTAACTTAATGAGCTTTAATGATGCTTCAAGAGCAAT 2220  
 QY 1315 SerLeuSerGluThrValLysLysValAlaTyrAspHisGlyLysLeuAsnLeuTyrLeu 1334  
 Db 2221 TCTCTGACGAGACAGTGAAGAGTGAATGATCATGGAATAACTCAACCTCATATTA 2280  
 QY 1335 AspSerValAsnGluThrGlnPheCysValAsnIleProAlaValArgAsnPheLysVal 1354  
 Db 2281 GATTCGTGAATGAAGAACCAAGTTTGTTGTAATATCTGCTGTGAGAACTTTAAAGTT 2340  
 QY 1355 SerAsnThrGlnAspAlaSerValSerIleValAspTyrTyrGluProArgArgGlnAla 1374  
 Db 2341 TCAATTAACCAAGATGCTTCAAGTGTCCATATGATGATTAATGAGCCAAAGGAGACAGCG 2400  
 QY 1375 ValArgSerTyrAsnSerGluValLysLeuSerSerCysAspLeuCysSerAspValGln 1394  
 Db 2401 GTGAGAAATTTCAACTCTGAAGTGAAGTGTCTCTCTGTGACCTTTGAGAGATGTCCAG 2460  
 QY 1395 GlyCysArgProCysGluAspGlyAlaSerGlySerHisHisHisSerSerValIlePhe 1414  
 Db 2461 GGCCTCGCTCTTGTGAGAGATGAGACTTCAAGCTTCCATCATCACTCTTCACTATTTT 2520  
 QY 1415 IlePheCysPheLysLeuLeuTyrPheMetGluLeuTrpLeu 1428

Db 2521 ATTTCTGTTCAAGCTTCTGACTTATGGAACCTTG3CTG 2562

RESULT 13  
AK095888  
LOCUS AK095888 2273 bp mRNA linear PRI 15-JUL-2002  
DEFINITION Homo sapiens cDNA FLJ38569 f18, clone HCHON2006459.  
ACCESSION AK095888  
VERSION AK095888.1 GI:21755236  
KEYWORDS oligo capping; f18 (full insert sequence).  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
AUTHORS  
1  
Oshima, A., Takahashi-Fujii, A., Tanase, T., Imose, N., Takeuchi, K., Arita, M., Mutsaers, K., Yuuki, H., Hara, H., Sugiyama, T., Irie, R., Otsubo, T., Sato, H., Makamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Matsushima, M., Murakawa, K., Kanehori, K., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahara, K., Masuno, Y., Nagai, K. and Isegaki, T.  
NEO human cDNA sequencing project  
Unpublished  
2 (bases 1 to 2273)  
Isegaki, T. and Yamamoto, J.  
Direct Submission  
Submitted (04-JUL-2002) Takao Isegaki, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan  
(E-mail: genom@hri.co.jp, tel: 81-438-52-3975, fax: 81-438-52-3986)  
NEO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.  
Location/Qualifiers  
1..2273  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="HCHON2006459"  
/cell\_type="chondrocytes (HC)"  
/clone\_lib="HCHON2"  
/note="Cloning vector: pME18FLJ3-Primary culture, chondrocytes"  
66..2063  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="AAC06442.1"  
/db\_xref="GI:21755237"  
/translation="MQGPPILTAALLCVCTAALAVAGPRFLVTAAGIIRPGNVTI  
GVEILHEPGQVTVKABELKTASNLTVSEAEVGFKEGSKTLTLPSPNSDELY  
ELRYGTQDEILFNSRSTSFETKRIASFVOTKALYKPEQVFPFVILFSPFKY  
KTSNLIILDKSNLIQOMLSQSDIGVISTFOLSPHILIGMSIOVONDDVYOS  
FOVEYVLPKREYVLTQPLPYCSMNSKHLNGITTKATYGRVKDVTILPLSPMK  
KKNITKTKFKNLSANFSFNDENKNDMSNGSEYVLDLSPGVEILTVTESVTGI  
SRNVTNFFKQHDYIIEFPDYTVLKPSLNTATVATVTRADGQTLERNNVIT  
VTORNATYEMSGNSGNDKMEAVQIKNTVYPOSGTPEKIEPILDESSELKAFILS  
KSNVAHSLFKSPKTYIOLKTRDENIKVSGPELVVSGNRLSELSYMVSRQLVA  
VGNKSTMFSLTPENSWTPKACVIVYIEDDELISPVAKIPQVLVKNKIKLYMSKY  
KAERSEKSLISTVQPSIYGIYAVDKSVLMAASNDITMENVHLELYNTGYLIG  
MFMNSFNVFGCGLMVLDANLTYDVIDGVNLFQTEAL"

BASE COUNT 712 a 431 c 462 g 668 t  
ORIGIN

Alignment Scores:  
Pred. No.: 6.9e-212 Length: 2273  
Score: 3360.00 Matches: 657  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 45.73% Indels: 0  
DB: 9 Gaps: 0  
US-10-020-095-4 (1-1428) x AK095888 (1-2273)

Qy 1 MercGlnGlyProProLeuLeuThrAlaAlaHisLeuLeuCyValCyThrAlaAlaLeu 20  
Db 66 ATGAGAGCCCAACCGCTCCGACCGCGCCGACCTCTCTGCGTGTGCACGGCGCGCTG 125  
Qy 21 AlavAlaLapProGlyProAArgPheLeuValThrAlaProGlyIleIleAArgProGly 40  
Db 126 GCGTGGCTCCGGGCTCGGCTTCTGTTGTGTACAGCCCAAGGATCATCAGCCCGAGGA 185  
Qy 41 AsnValThrIleGlyValGluLeuLeuGluHisCyProSerGlnValThrValAla 60  
Db 186 AATGTACTATTGGGGGAGGAGCTTCGGAGACATGCGCTTCACAGGTGATGGAAGCG 245  
Qy 61 GluLeuLeuLysThrAlaSerAsnLeuThrValSerValLeuGluValGluGlyValPhe 80  
Db 246 GAGTGTCTCAAGACAGCATCAAACTCAGTCTGTCTGTCTGAGAGCAGAGAGTCTTT 305  
Qy 81 GluLysGlySerPheLeuThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGlu 100  
Db 306 GAAAGAGCTCTTTTAAGACACTTACTCTTCATCACTACTCTGAAACAGTCAGATGAG 365  
Qy 101 IleTyrGluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThr 120  
Db 366 ATTTATGAGTACGTGTATACCGGACGTCACAGATGAGATTTATTTCTTAATATGATAC 425  
Qy 121 ArgLeuSerPheGluThrLysArgIleSerValPheIleGlnThrAspLysAlaLeuTyr 140  
Db 426 CGCTTATCATTTAGACCAAGACAAATATCTGCTTCAATCAACAGCAAGCGCTTATAC 485  
Qy 141 LysProLysGlnGluValLysPheArgIleValThrLeuPheSerAspPheLysProTyr 160  
Db 486 AAGCAAGAGAGAGAGGAGTTTCGATTTGATCACTCTTCAGATTTTAAGCTTAC 545  
Qy 161 LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGlnTyrLeu 180  
Db 546 AAAACCTCTTAAACNTCTCATTTAAGAGCCCAATCAATTTGATCCAAAGTGTG 605  
Qy 181 SerGlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIle 200  
Db 606 TCACAACAAAGATGATCTTGAGTCATTTCCAAACCTTTCAAGCATCTTCCCATCAATA 665  
Qy 201 LeuGlyAspTrpSerIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGln 220  
Db 666 CTGTGACTGCTCTATTCAGTTCAAGTGAATGACCAACATCAATCAATCAATTCAG 725  
Qy 221 ValSerGluTyrValIleuProLysPheGluValThrLeuGlnThrProLeuTyrCySer 240  
Db 726 GTTTCGAATATGATATTACCAAAATTTGAAGTACCTTTCAGACACACATTATTTGTC 785  
Qy 241 MetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProVal 260  
Db 786 ATGAATCTTAAGCATTTAATATGATACATCAAGCAAGATATCATATGGAAGCAGTG 845  
Qy 261 LysGlyAspValThrLeuThrPheLeuProLeuSerPheTrpGlyLysLysLysAsnIle 280  
Db 846 AAGGAGACCTAAGCTTAATTTTACCTTTTGGGGAAGAAAGAAATAT 905  
Qy 281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluGluMetLys 300  
Db 906 ACAAAAACATTTAAGATTAATGATGATTCGCAAACTTCTTTTAATGATGAAGATGAAA 965  
Qy 301 AsnValMetAspSerSerAsnGlyLysSerGluTyrLeuAspLeuSerSerProGlyPro 320  
Db 966 AATGTATGATCTTCAATGATGATCTTTCGATATCTGATATCTTATCTTCCCTGAGCA 1025  
Qy 321 ValGluIleLeuThrThrValThrGluSerValThrGlyLysSerArgAsnValSerThr 340  
Db 1026 GTAGAAATTTTAAACAAGTGAAGATGATGATGATTTTCAAGAAATGTAAGCACT 1085

```

QY      341 AsnValPhePheLeuGlnHISaPylrIleIleGluPhePheAspYrThrValLeu 360
DB      1086 AATGTTCTTCAAGAAACATGATTCATCATGAGTTTTCATTAATCTACTGCTTG 1145
QY      361 LysProSerLeuAsnPheThrAlaThrValValThrArgAlaAspGlyAsnGlnLeu 380
DB      1146 AACGCTCTCTCAACTTCACAGCCACTGGAAGGATGCTGCTGATGGCAACACAG 1205
QY      381 ThrLeuGluGluArgArgAsnValValIleThrValThrGlnArgAsnYrThrGlu 400
DB      1206 ACTCTGGAAGAAAGAAATATGATGATCAATACAGTGCACAGAGAACTTACTGAG 1265
QY      401 TyrTrpSerGlySerAsnSerGlyAsnGlnLysMetGluAlaValGlnLysIleAsnYr 420
DB      1266 TACTGAGCGGACTTAACGTGAGAAATCAGAAATGAGAACTGTCAGAAATTAATAT 1325
QY      421 ThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGlnAspSerSerGlu 440
DB      1326 ACTGTCCTCCAAAGTGAACCTTTTAAGATTGAATTCCTCAATCTGAGAGATTCAGTGA 1385
QY      441 LeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSerLeuPhe 460
DB      1386 CTACAGTTGAGGCTTATTTCTTGGTAGTAAAGTACAGTGCAGTTCATGATCTGTT 1445
QY      461 LysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGlnAsnIleLysValGly 480
DB      1446 AAGTCTCTAGTAAAGCATCATCCAACTTAAACAGAGATGAAATATTAAGGTGGGA 1505
QY      481 SerProPheGlnLeuValValSerGlyAsnLysArgLeuLysGlnLeuSerYrMetVal 500
DB      1506 TCCTCTTTGAGTGTGGTGTAGTGCACAAACGATGAGAGATTAGCTATATAGTA 1565
QY      501 ValSerArgGlyGlnLeuValAlaValGlyLysGlnAsnSerThrMetPheSerLeuThr 520
DB      1566 GTATCCAGGGGACAGTTGGTGGCTGAGAGAAACAAATTCACAAATGCTCTTTAA 1625
QY      521 ProGlnAsnSerTrpThrProLysAlaCysValIleValTyrIleGlnAspAspGly 540
DB      1626 CCAGAAATATCTTGACCTCCAAAGCCTGTGTAATGTGATTAATTAAGATGATGG 1685
QY      541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560
DB      1686 GAAATATATAGTATGTTCTTAAATAATCTGTTCACTGTTTAAATAAGATTAAG 1745
QY      561 LeuTyrTrpSerLysValLysAlaGluProSerGlnLysValSerLeuArgIleSerVal 580
DB      1746 CTATATTGAGTAAGTGAAGCTGAACCATCTGAGAAAGTCTCTTAGATCTGTG 1805
QY      581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600
DB      1806 ACACAGCTCGACTCCATAGTGGATTTGAGCTGTTGACAAAGTGTGAATCTGATGAT 1865
QY      601 AlaSerAsnAspIleThrMetGluAsnValValHisGlnLeuGlnLeuLysTrpAsnThrGly 620
DB      1866 GCCTCAATGATTTTCAATGAGAAATGTGTCATGAGTGGAACTTTATTAACACAGA 1925
QY      621 TyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGlnLysGlyLeuTrpVal 640
DB      1926 TATTAATTAGGCAATGTTCAATGAAATCTTTGACGCTTTTCAGAAATGAGACTCTGG 1985
QY      641 LeuThrAspAlaAsnLeuThrLysAspYrIleAspGlyValTyrAspAsn 657
DB      1986 TTGACAGATGCAAACTCAAGAAAGATTAATTAATGAGTGTATTATACAAAT 2036

```

```

RESULT 14
AX054953      4527 bp      DNA      linear      PAT 13-JAN-2001
LOCUS      AX054953
DEFINITION      Sequence 68 from Patent WO0073328.
ACCESSION      AX054953
VERSION      AX054953.1 GI:12228337
KEYWORDS
SOURCE      Caenorhabditis elegans
ORGANISM      Caenorhabditis elegans

```

```

REFERENCE
AUTHORS      Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditidae;
TITLE      Rhabditidae; Rhabditidae; Peloderinae; Caenorhabditis.
JOURNAL      1
            Van Crieke, W., Roelens, I., Bogaert, T. and Verweide, P.
            Unc-5 constructs and screening methods
            Patent: WO 0073328-A 68 07-DEC-2000;
            Devgen NV (BE)
FEATURES
source
            location/Qualifiers
            /organism="Caenorhabditis elegans"
            /mol_type="genomic DNA"
            /db_xref="caxon:6239"
BASE COUNT      1313 a      948 c      1077 g      1189 t
ORIGIN
Alignment Scores:
Pred. No.:      2,276-112      Length:      4527
Score:      1854.50      Matches:      476
Percent Similarity:      51.23%      Conservative:      292
Best Local Similarity:      31.75%      Mismatches:      562
Query Match:      25.24%      Indels:      169
DB:      6      Gaps:      44
US-10-020-095-4 (1-1428) x AX054953 (1-4527)
QY      15 ValCysThrAlaAlaLeuAlaValAlaProGlyProArgPheLeuValThrAlaProGly 34
DB      82 GTGTCAACACCGCGCGCCAGTTAAACACGCG---ACTTATATGCTTGTGCCCCACCA 138
QY      35 IleIleArgProGlyGlyAsnValThrIleGlyValGlnLeuLysGlnHisCysProSer 54
DB      139 GTAAGTGCACCCGATCAACCGCTTCCGTTTGCATGACGACCTTGAAACAGACTCGAT 198
QY      55 GlnValThrValLysAlaGlnLeuLysThrAlaSerAsnLeuThrValSerValLeu 74
DB      199 GAAAGCATGATTTGCGATGGAAGTGAACCTGAACGGAACGAGACGATTCGCTCGCA 258
QY      75 GluAlaGluGlyValPheGlnLysGlySerPheLysThrLeuThrLeuProSerLeuPro 94
DB      259 GTTATAGCAATTTG---AAGCGGGAATCCGACGAGCTGTGAGCTTAAGTAGATGCT 315
QY      95 LeuAsnSerAlaAsp-----GluIleTyrGlnLeuValGlnValThrGlyArgThrGlnAsp 112
DB      316 GCTCATCATCTAATCTCTAAGCAAGCTATAGCTCTACATCCCGGGGAGACTCTTAAC 375
QY      113 ---GluIleLeuPheSerAsnSerThrArgLeuSerPheGlnThrLysArgIleSerVal 131
DB      376 GCAGAACTCATATTTGAAACGAAATGAGTTAAAGTACGATCAAAAAGCTTTATCAGTT 435
QY      132 PheIleGlnThrAspLysAlaLeuTyrLysProLysGlnGluValLysPheArgIleVal 151
DB      436 TTATCTCAACACGATCGAGCAATCTATGCGCTGCTTCTTGTTCGATACCGGGCAATT 495
QY      152 ThrLeuPheSerAspPheLysProTyrLysThrSerLeuAsnIleLeuIleLysAspPro 171
DB      496 GTAAGTAATCGAATCTTAACACCTTACGCGAAGACGTACTATTAAATTTTGCATGCA 555
QY      172 LysSerAsnLeuIleGlnGlnThrLysSerGlnLysAspLeuGlyValIleSerLys 191
DB      556 AGCAGAAATCTGATTTCAACAGACGATGCTGTAACCTTGATAGAGAGATTAACAGCGCG 615
QY      192 ThrPheGlnLeuSerSerHisProIleLeuGlyAspTrpSerIleGlnValGln----- 209
DB      616 GAGCTCCAACTTGCGGAAGAGACTCTCTTGGAGATGCTTTATTCAGAGGAAACCTCG 675
QY      210 -----ValAsnAspGlnThrTyrTrpGlnSerPheGlnValSerGluTyrValLeuPro 227
DB      676 AATGAGTCCAGAGAAATCG-----TCATTACGTAAGTAACTACGCTTCTCC 726
QY      228 LysPheGluValThrLeuGlnThrProLeuTyrCysSerMetAsnSerLysHisLeuAsn 247
DB      727 AAATTTGAAGTAATATCAAGACGCTTCGTTATTAATTAACGAT---GATTGTGCA 783

```

OY	248	GLYTHRIETHRALALYSERYTHRTRYGLYSPROVALYSLYASPVALTHTREUTHR	267
DB	784	GTCTTCGTGGATGCAAGATGACATATGTGAAAAAGCGTTTGTGTAAGAGCCAAAGTATCT	843
OY	268	PHE---LEUPROLEUSERPHEPTRP-----	275
DB	844	CTAGAGCTTCCATGGCACAGATGGCATGGATGTTCCCATATCATTTGACGAAATATGGA	903
OY	276	LYSLYLYSASPN-----LLETHRYSTHRYPHELYSILEASN-----GLYSER	289
DB	904	GTCAGAAAGAGAGAGAGCTGATGCTCGAGCGCTACTGTAAAGCTGAACAGAACAGAGAA	963
OY	290	ALAASNPHESERPHESASNPOLUICULIWEIYSASNVALMETASPSERASNDIYEU	309
DB	964	GGGCGAGTGTGTTTCTTCGATGATATAGCTCAAGAGCG-----	999
OY	310	SERGIUPLYRLEUASPSEUSERPROGLYPROVALGIULEUETHRVALTHRGU	329
DB	1000	CACAACTGCTACATGAAATGGGAGAGAGTTCAATTCCGATTTGTCCTCACTGTA	1055
OY	330	SERVALTHRGYLISSERARGHNSALSERTHASNVALPHEPHELYSGLNHISAPTRY	349
DB	1060	GACATCACTGAATATGAAAGAAAGCACTACTCATCAAAATTTCACGTCGTAAGAGATGC	1119
OY	350	ILEILEGLIUPHEPHEAPTYYRTHRVALLEULYSPROSERLEUASNPHETHALATHR	369
DB	1120	AAACTGTAGTGTGAAAGAGAGAGACACCTTCAAGCCAGACACTCATTTATATGTTGTA	1179
OY	370	VALLYSVALITHRARGALASPGLYASN-----GLNLEUTHRLEUGLUNHARGASPN	387
DB	1180	GTTGCACCTGAAACAAATGATGATACACTCCCGTAAGGCACTTGTCCAAAGAGA----	1233
OY	388	ASNVALVALIETHRVALTHRGINARGASNTRYTHRGUPLYRTRIPSEGLYSERASNSE	407
DB	1234	-----GTTACAGATATCAACATTTACACATATCA-----TACAAATCATGACACTTCA	1281
OY	408	GLYASNGINLYSMETGLUVALGLNLYSILEASNTRYTHRVALPROGLINSERGLYTHR	427
DB	1282	TCTCTGACAAAGAGAAACAAACAAATTTGGGAAGTTGACGCTACAGCGACCTCGGTT	1341
OY	428	PHELYSILEGLIUPHEPROILEUENLUNASPSERSEGLULEUGLLEULYALATYRPH	447
DB	1342	CTCACCCCTCCACCCACCAATCACTCACTTGTGCTCGCATTTGAGGCTCAT---TATCAT	1398
OY	448	LEUGLYSERLYSSERSEMETALVALHISERILEUPHELYS-----	461
DB	1399	ATTGTTGGCAAGACACACTTCCACGCCACCCCAATATATTCCTCGCTATATGTGAGGCA	1458
OY	462	-----SERPROSERLYSTHTRYRILEGULEULYSRTHARGASPGUNSN--ILELYS	478
DB	1459	GCCGTGACCCCAACAGAGATTTCTCGACAGCTTTTGGCTGATAAAGAGAGAGCTGTTAT	1518
OY	479	VALGLYSERPROPHGLULEUVALSERGLYASNLYSARGLEULYGLULEUSERTRYR	498
DB	1519	GTAGGAAAGAGCTTAACTCTTCGTGAAGGCTACCCAGCGCTGTGCACATTTACAT	1578
OY	499	METVALVALSERARGLYGULEUVAL-----ALAVAGLYLSEGLASN	513
DB	1579	CAGGTGATGTCTCGTTGCAACATTTGTGGTCTCAACAATGACGTATATTTCCGAACAT	1638
OY	514	SERTHMETPHESETRLEUTHRPROGLUNASERTPRNRPOLYSEALCYSEVALILEVAL	533
DB	1639	GCGACACATT---AGCTTCCACAGCAAGCGTAAATGAGCCCAAAATTCAGGTGATGCTGC	1699
OY	534	TYR---TYRILEGLIUNASPSGLIULILELLESERASPVALLEULYSILEPROVALGIN	552
DB	1696	TATGCCATTATTTAGAGATGATCAAAAGATCTTGTGTATGACCTTGACTTTAAAGTAAA	1755
OY	553	LEUVALPHELYSASNLYSILELYSEULYRTRIPSETRYVALYSLYALAGLIPROSEGLU	572
DB	1756	GGAATATTTCCAAACCAAGTGGCACTTATTCATTTGACAAACAGACAGTGAAGCCAGAGAG	1815
OY	573	LYSVALSERLEUARGILESERVALTHRGINPROASERILEVALGILEVALALVAL	592

Db 1816 AATGGAAGCTTTAAGGTCACA---CTGACAAAGAACTTTTGTGTGGCCTCTCTTGCTA 1875  
 Oy AspYSerSerValAsnLeuMetAsnAlaSerAsnAspIleThrMetGluAsnValAlaHis 612  
 593 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 612  
 Db GATCAGAGTGTCTCTCTGTATAAAACCTGGCAATGATATCACTAGAGAAAAGTTGAGCAA 1933  
 Oy 613 GluLeuGluLeuTyraSerThrGlyTyrTyrrLeuGlyMetPhe----- 626  
 613 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 626  
 Db GACTTGGAGAAATTATGACTCTTAATATGTTGGAGAGAGGATTCGGGGGTCCACGGCATGG 1992  
 Oy 627 -----Met 627  
 Db 1993 GAGCCAATTGATAGAAAGAAAGAGCTCCATTGGAGACCTTGCTGGGGAATTGGAGGAGGT 2055  
 Oy AsnSerPheAlaValPheGlnGluCysGlyLeuTrpValLeuThrAspAlaAsnLeuThr 647  
 628 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 647  
 Db GATGCTCAATCAATCTTTTCTTAAGCTGGGCTTGAGTTTGAACCGACGCCCTCTTGTAAC 2111  
 Oy 648 LysAsp-----TyrIleAspGlyValTyrAspAsnAla 658  
 648 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 658  
 Db 2113 CGCGGAGCTCAACAGTAATTATATGCTACGTAATGATGATGAGATGGCTCCAGAAATGGCT 2172  
 Oy GluTyrAlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPhe 678  
 659 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 678  
 Db 2173 GAAGCCGGA-----TTGCCGCTCCGCCAATGGGC----- 2207  
 Oy 679 SerLeuGlySerSer-----ProHisValArgLysHisPheProGluThrTrpIle 695  
 2203 -----GGGTGCTGGCCTCCACCGCCGACAGTTGGAAGTTCTTCCACACACTCGGATT 2255  
 Db 696 TrpLeuAspThrAsnMetGlyTyrArgIleTyrGlnGluPheGluValThrValProAsp 715  
 2257 TGGTGGGATCTCAAC---TCTACTAGTGGAGAAAGTTGAGATGGAATGGAAGCTCCGGAC 2311  
 Oy 716 SerIleThrSerTrpValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeu 735  
 2314 ACCATCACTCTGTGGTGTGCTCTCCACTTTCGAATCAATGAGAGAAATGGTCTCGGTGG 2377  
 Db 736 ThrThrThrProValGluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuPro 755  
 2374 GCTCCACAGACTTCAAAACTTCGCGTGTTCGTCATTTTATTTCACTCACTCATCTTCCA 2433  
 Oy 756 TyrSerValIleArgGlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLys 775  
 2434 TATGCTGTCGCTGGGAGAAAGTTGGCACTTTTGGCTGTGTTTCAATTAATATGAGA 2493  
 Db 776 AspAlaThrGluValLysValIleIle-----GluLysSerAspLysPheAspIleLeu 793  
 2494 AAGAGCAAGATGTAAAGTACTCTGAAATGATATTAAGACTGCTGTAGCATCTCTCG 2555  
 Oy 794 MetThrSerSerGluIle-----AsnAlaThrGlyHisGlnInThrLeuLeuVal 810  
 2554 AAGAAAGATGGAACCTGATAGGACGCGTACAGAGGTGGCCAAACAAATGTTCCGATTGTG 2613  
 Db 811 ProSerGluAspGlyAlaThr-----ValLeuPheProIleArgProThrHisLeu 827  
 2614 TCCGTGTGCTGGTGGAGAAATCGAAGACAGATGTACTTCCCAATGTGTCCTTCATATT 2672  
 Oy 828 GlyGluIleProIleThrValThrAlaLeuSerProThrAlaSerAspAlaValThrGln 847  
 2674 GGAGAGATCCCGGTACATATAGGCGTATTCATCCCAAGGTGGTGAAGTACATAGATG 2733  
 Db 848 MetIleLeuValLysAlaGluGlyIleGluLysSerTyrSerGlnSerIleLeuLeuAsp 867  
 2734 AACTTGACAGTGTATCCCAAGAGATATTAAGATAGATAAAATATTCCTATTGTCTATGAC 2799  
 Oy 868 LeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeuSerPheSerPheProProAsn 887  
 2794 CTCAACATATACCTCTGGATTTCTCC---AAAATCTTGAGCTTATCTGGCCAAAGAC 2855  
 Db 888 ThrValThrGlySerGluArgValGlnIleThrAlaIleGlyAspValLeuGlyProSer 907









QY 685 HisValArgLysHisPheProGluThrTrpPheLeuAspThrAsnMetGlyTyrArg 704  
 DB |||||  
 DB 2257 ACAGTTCGAAAGTTCCTCCACACACTGGATTGGCTCGACTCACTACTACT 2313  
 QY 705 IleTyrGlnGluPheGluValThrValProAspSerIleThrSerTrpValAlaThrGly 724  
 DB |||||  
 DB 2314 GGAGAGATTGAGATGAAATTGAAAGTCCGACACATCACTCGTGGTCCGCTTCACT 2373  
 QY 725 PheValIleSerGluAspLeuGlyLeuGlyLeuThrThrProValGluLeuGlnAla 744  
 DB |||||  
 DB 2374 TTTCGACATCATAGAGAAATGCTCGGTGGTCCCAACGACTCAAAACTGCGCG 2433  
 QY 745 PheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArgGlyGluLeuPhe 764  
 DB |||||  
 DB 2434 TTTCGACATCATAGAGAAATGCTCGGTGGTCCCAACGACTCAAAACTGCGCG 2493  
 QY 765 AlaLeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrGluValValIleIle 784  
 DB |||||  
 DB 2494 GCACCTTTGGTCTGTTTCAATTATATGAAAGAGCAAGATGTAACAGTACTG 2553  
 QY 785 -----GluLysSerAspLysPheAspIleLeuMetThrSerSerGluIle----- 799  
 DB |||||  
 DB 2554 AAGTATGATAAAGACTCTGTATACATCTCTGAGAGAAAGATGAACTGTATAGGCGT 2613  
 QY 800 AsnAlaThrGlyHisGlnGlnThrLeuLeuValProSerGluAspGlyAlaThr----- 817  
 DB |||||  
 DB 2614 GACGAGTTGGCCCAAAATGCTCGATTGGTCCGTTGCTGAGAGAAATGCAATGCA 2673  
 QY 818 ---ValLeuPheProIleArgProThrHisLeuGlyGluIleProIleThrValThrAla 836  
 DB |||||  
 DB 2674 GCAGTGTACTCCCAATGTGCGTCTCAATTGAGAGATCCCGTATCAATTAAGCGT 2733  
 QY 837 LeuSerProThrAlaSerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIle 856  
 DB |||||  
 DB 2734 ATTCGATCCCAAGTGTGTATGACAGAGATGAACTTGAAGTTCATCAACAAGATAT 2793  
 QY 857 GluLysSerTyrSerGlnSerIleLeuAspLeuThrAspAsnArgLeuGlnSerThr 876  
 DB |||||  
 DB 2794 AAGAGATATAGAAATATTCATTCATTCATTCATTCATTCATTCATTCATTC 2853  
 QY 877 LeuLysThrLeuSerPheSerPheProProAsnThrValThrGlySerGluArgValGln 896  
 DB |||||  
 DB 2854 ---AAAAATCTTGAGCTTATCTGCGCCAAAGCAGCTGTGATGATGCGAGAGCA 2910  
 QY 897 IleThrAlaIleGlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerIleArg 916  
 DB |||||  
 DB 2911 TTAGATGATTTGGTATGATGATGAGACCAAGTTCATTAACATGCTCATAGCTG 2970  
 QY 917 MetProTyrGlyCysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeu 936  
 DB |||||  
 DB 2971 ATGCCATATGATGATGAGACCAAGATATGATGATGATGATGATGATGATGATG 3030  
 QY 937 AspTyrLeuThrLysLysLysLeuThrAspAsnLeuLysGluLysAlaLeuSerPhe 956  
 DB |||||  
 DB 3031 AAGTCTTAAAGCAAAATATGAAATGAGACCAATTAAGCAAAAGCTTCAAGTTT 3090  
 QY 957 MetArgGlnGlyTyrGlnArgLeuLeuLysGlnArgGluAspGlySerPheSerAla 976  
 DB |||||  
 DB 3091 ATTGACCAAGGATTCAGAGAGATGAGACCAAGCTGCAAAAGCTGCAAAAGCT 3150  
 QY 977 PheGluAsnTyrAspProSerGlySerThrTyrLeuSerAlaPheValLeuArgCysPhe 996  
 DB |||||  
 DB 3151 TTTCGACATCATAGAGAAATGCTCGGTGGTCCCAACGACTCAAAACTGCGCG 3210  
 QY 997 LeuGluAlaAspProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTrp 1016  
 DB |||||  
 DB 3211 CATCAAGCCAAACATATCCATTCGATCTCAATGATATTCGCGAGACATCCCATC 3270  
 QY 1017 LeuLysGluHisGlnLysSerAsnGlyLysPheTrpAspProGlyArgValIleHisSer 1036  
 DB |||||  
 DB 3271 TTGAATCTCAAAATGAGATGCTGTGATGATGATGATGATGATGATGATGATGAT 3330  
 QY 1037 GluLeuGlnGlyGly---AsnLysSerProValThrLeuThrAlaTyrIleValThrSer 1055

DB |||||  
 DB 3331 GATATGAGGAGGAGTCTCAGACCGGGGTGAGCTTCAAGCGCATTTGTGTATTCG 3390  
 QY 1056 LeuLeuGlyTyrArgLysTyrGlnProAsnIleAspValGlnGluSerIleHisPheLeu 1075  
 DB |||||  
 DB 3391 ATTTTG-----GAGATGAGATGAGAAATGAGAAAGCGGTGACGATTTTG 3435  
 QY 1076 GluSerGluPheSerArgGlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyrAla 1095  
 DB |||||  
 DB 3436 GAAAGCATTTGATGAGATGATGAGAAATGCGTATCAATGCGTGTAGTGCATATGCA 3495  
 QY 1096 LeuSerSerValGlySerProLysAlaLysGluAlaLysAsnMetLeuThrTrpArg--- 1114  
 DB |||||  
 DB 3496 CTTCAATTGGCAAAAAGTACCAAGCTGAAAGGATTCGAGATTTGAGAAACATTAAG 3555  
 QY 1115 ---AlaGlnGlnGlyGlyMetGlnPheTrpValSerSerGlu-----SerLysLeu 1131  
 DB |||||  
 DB 3556 ATCGTTGAGAAAGTGGAGATGAGATTTGCGCACTGCTCAGAAAAAGTGAAGAACTA 3615  
 QY 1132 SerAspSer-----TrpGlnProArgSerLeuAspIleGluValAlaAlaTyr 1147  
 DB |||||  
 DB 3616 AAGATATGAGAGCGTATATGTTCCAGGCCGTCAGTATGACATTTGAACATACCTTAC 3675  
 QY 1148 AlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGluGlyIleProIleMetArgTrp 1167  
 DB |||||  
 DB 3676 GCAGTCTTTCTTACCTGCGCCCAAAATCAAACTCAGAACTCTCTCAATTATTCGTTGG 3735  
 QY 1168 LeuSerArgGlnArgAsnSerLeuGlyLysPheAlaSerThrGlnAspThrThrValAla 1187  
 DB |||||  
 DB 3736 TTAGATTCCAACGCAAGAGCTAGAGAGATTCCTTCACCTCAGACACATGCTGCGT 3795  
 QY 1188 LeuLysAlaLeuSerGluPheAlaAlaLeuMetAsnThrGlnArgThrAsnIleGlnVal 1207  
 DB |||||  
 DB 3796 CTTCAAGCTCTCTCTCTTACCGACAGCTCACTTATTCGACAAACACAAAGTAACTA 3855  
 QY 1208 ThrVal---ThrGlyProSerSerProSer-----ProLeuAlaVal 1220  
 DB |||||  
 DB 3856 ACAATTCACAGCAAAACATATCTACATTCATTCATTCATTCATTCATTCATTC 3915  
 QY 1221 ValGlnProMet-----AlaValAsnIleSerAlaAsnGlyPhe 1233  
 DB |||||  
 DB 3916 CTTCAATCTTATCAACTATTCATTCATTCATTCATTCATTCATTCATTCATTC 3975  
 QY 1234 GlyPheAlaIleCysGlnLeuAsnValValTyrAsnValLysAlaSerGlySerSerArg 1253  
 DB |||||  
 DB 3976 GGTGGTCTTCGCTCACTTATGATTCCTTAC----- 4008  
 QY 1254 ArgArgArgSerIleGlnAsnGlnGluAlaPheAspLeuAspValAlaValLysGluAsn 1273  
 DB |||||  
 DB 4009 TACAGGACTCTTTGAATGAGATGCTCGTTCCTTCGCTCCAGGAGATCAAGGAATTA 4068  
 QY 1274 LysAspAspLeuAsnHisValAspLeuAsnValCysThrSerPheSerGlyProGlyArg 1293  
 DB |||||  
 DB 4069 AAGCTGTGA---AATCGCATGCAATTCGATTTGTTGTTGTTGTTGTTGTTGTTG 4125  
 QY 1294 SerGlyMetAlaLeuMetGluValAsnLeuSerGlyPheMetValProSerGluAla 1313  
 DB |||||  
 DB 4126 TTTAATCATGCTTTAGCCGAATGATGATGCTTGTAGTGTATCCGATTCGATGCGGAA 4185  
 QY 1314 IleSerLeuSerGluThrValLys-----LysValGluTyrAspHisGly 1328  
 DB |||||  
 DB 4186 GTGACTACTTTGACAAAGATTTGAGATTTGCAAAAGGAGAGATGAAAGACACACT 4245  
 QY 1329 LysLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCysValAsnIleProAla 1348  
 DB |||||  
 DB 4246 AAGATGAACGTATATTTTCAACCGCTCGGTGCGCTCGGTATGCTCTCGTTATACCT 4305  
 QY 1349 ValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSerIleValAspTyrTrp 1368  
 DB |||||  
 DB 4306 GATGCTACTTATCAAGTTGCGATCAAAACCTGCAATTCGATGATGATGATGATGAT 4365  
 QY 1369 GluProArgGlnAlaValArgSerTyrAsnSerGluValLysLeuSerSerCysAsp 1388  
 DB |||||

Db 4366 GATCCGAGAGACACACTCAAAAATGCAATACGCCGCAAGCAAAACGATCGCTGCAGAG 4425

Qy 1389 LeuCySerAepValGlnGlyCyAargPro 1398  
|||  
::: ||| |||

Db 4426 AAATGTGGG-----GAAGACTGTGTGGCCA 4449

Search completed: January 17, 2004, 08:32:16  
Job time : 11145 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: January 16, 2004, 20:57:28 ; Search time 753 Seconds

(without alignments)  
5119.253 Million cell updates/sec

Title: US-10-020-095-4

Perfect score: 7348

Sequence: 1 MGGPPLTFAHDLVCVTAL.....HSSVIFPFCFKLYFMEIWL 1428

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2552756 segs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seg length: 0  
Maximum DB seg length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DBV=xld  
-Q=/cgn2.1/USPro.spool\_p/US10020095/rnuc.15012004.163821.18701/app\_query.fasta\_1.1607  
-DB=N.Geneseq.150un03 -QFMT=fastag -SUFFIX=ring -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10020095 @CCN 1.1 603 @runat.15012004.163821.18701 -NCPU=6 -ICPU=3  
-NO\_MAP -LARGEOUTERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DBV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FCAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N.Geneseq.19Jun03:\*

1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*

2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*

3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*

4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*

5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*

6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*

7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*

8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*

9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*

10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*

11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*

12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*

13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*

14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*

15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*

16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*

17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*

18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*

19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*

20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*

21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*

22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*

23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*

24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	7331	99.8	4369	25	AAD49434	Human r150 gene #1
2	7326.5	99.7	4335	24	ABL49815	Human platelet all
3	7326.5	99.7	4761	24	ABO79964	Human CD109 K1 pro
4	7326.5	99.7	5882	25	AAD49440	Human blood cell s
5	7326.5	99.7	5895	24	ABO79966	Human CD109 K1-H7
6	7318.5	99.6	5883	25	AAD49435	Human r150 DNA #1.
7	7317.5	99.6	4335	24	ABL49816	Human platelet all
8	7317.5	99.6	4761	24	ABO79965	Human CD109 K1 var
9	7317.5	99.6	5895	24	ABO79967	Human CD109 K1-H7
10	7307.5	99.4	4335	21	AAA60199	Hydrophobic domain
11	7307.5	99.4	4473	21	AAA62010	Hydrophobic domain
12	7081	96.4	4146	25	AAD49436	Human blood cell s
13	7068.5	96.2	4197	25	AAD49437	Human r150 DNA #2.
14	5470	74.4	3535	24	ABO79968	Human CD109 K15 pr
15	5461	74.3	3535	24	ABO79969	Human CD109 K15 va
16	1854.5	25.2	4527	22	AAC90951	C. elegans alpha-2
17	1850	25.2	4560	22	AAC90952	C. elegans alpha-2
18	1819.5	24.8	1448	24	ABL90478	Human polynucleoti
19	1767.5	24.1	1459	22	ABA08825	Human secreted pro
20	1533	20.9	5816	21	AAC74484	Human ORFX ORF3 p
21	1461.5	19.9	2403	24	ABV77330	Alpha-1 proteinase
22	1441	19.6	4577	21	AAA52113	Human alpha-2-macr
23	1440	19.6	4425	22	AAC90944	Human alpha-2-macr
24	1440	19.6	4577	22	AAD07359	Human DNA encoding
25	1440	19.6	4577	24	AAD45346	Human anti-microbi
26	1440	19.6	4577	24	ABN95738	Gene #2236 used to
27	1440	19.6	4577	24	ABK64513	Human benign prost
28	1440	19.6	4577	24	ABK12952	DNA encoding human
29	1440	19.6	4577	24	ABK24095	Human alpha2 macro
30	1440	19.6	4578	25	ABX74411	Human cDNA sequenc
31	1440	19.6	4578	25	ABX74412	Human cDNA sequenc
32	1440	19.6	4809	24	AAS94923	Human DNA sequence
33	1438	19.6	4527	22	AAH74529	A human alpha-2 ma
34	1438	19.6	4527	22	AAH74530	A human alpha-2 ma
35	1430	19.5	3612	23	ABL19705	Drosophila melanog
36	1428.5	19.4	4569	12	AAQ11176	Clone encoding rec
37	1425	19.4	4599	12	AAQ11581	Sequence encoding
38	1422	19.4	4488	24	ABSS5327	Human alpha-2-macr
39	1412.5	19.2	4595	24	ABK63624	Rat sequence diffe
40	1406	19.1	4278	23	ABL30183	Drosophila melanog
41	1401	19.1	4487	24	ABK92035	DNA encoding novel
42	1400	19.1	4615	23	ABN92193	DNA encoding novel
43	1400	19.1	4615	24	ABN92232	Gene #3730 used to
44	1376.5	18.7	4660	24	ABO93897	Human pregnancy zo
45	1376.5	18.7	6324	23	ABL19704	Drosophila melanog

#### ALIGNMENTS

RESULT 1

AAD49434

24-MAR-2003 (first entry)

Human r150 gene #1.

Human: GPI-anchored TGF-beta1 binding protein; r150 protein; cancer; glycosylphosphatidylinositol; transforming growth factor-beta1; therapy; gene; ds.

Human sapiens.

Key Location/Qualifiers

FT CDS 26..4312  
FT /tag= a  
FT /product= "Human r1520 protein #1"  
FT /transl\_except= (pos:2132..2134, aa:Xaa)  
FT /note= "Xaa corresponds to Ser, Tyr"  
XX  
XX WO200285942-A2.  
XX  
XX 31-OCT-2002.  
XX  
XX 24-APR-2002; 2002MO-CA00560.  
XX  
XX 24-APR-2001; 2001US-285713P.  
XX 14-FEB-2002; 2002US-356163P.  
XX  
XX (UYMC-) UNIV MCGILL.  
XX  
XX Philip A, Tam B;  
XX WPI; 2003-093100/08.  
XX P-PSDB; AAE32012.  
XX  
XX Novel transforming growth factor (TGF)-beta 1 binding reagent which  
XX comprises r150 protein which acts as accessory receptor of TGF-beta,  
XX useful for negatively modulating TGF-beta activity, and thus for  
XX treating cancer -  
XX  
XX Claim 7; Page 91-93; 127pp; English.  
XX  
XX The invention relates to novel transforming growth factor (TGF)-beta1  
XX binding reagent which comprises a glycosylphosphatidylinositol (GPI)-  
XX anchored TGF-beta1 binding protein referred to as r150 which acts as  
XX an accessory receptor of TGF-beta. The invention is used for negatively  
XX modulating TGF-beta activity, and thus for treating conditions  
XX characterised by overproduction of TGF-beta, such as cancer. Antisense  
XX molecules of the invention are used for increasing TGF-beta availability  
XX and increase graft success. The present sequence is a gene encoding  
XX human r150 protein.  
XX  
XX Sequence 4369 BP; 1301 A; 842 C; 932 G; 1293 T; 1 other;  
XX  
XX  
XX Alignment Scores:  
XX Pred. No.: 0 Length: 4369  
XX Score: 7331.00 Matches: 1425  
XX Percent Similarity: 99.86% Conservative: 1  
XX Best Local Similarity: 99.79% Mismatches: 2  
XX Query Match: 99.77% Indels: 0  
XX DB: 25 Gaps: 0  
XX  
XX US-10-020-095-4 (1-1428) x AAD49434 (1-4369)  
XX  
XX 1 MetGInGlyProProLeuLeuThrAlaAlaHisLeuLeuGlyValGlyThrAlaLeu 20  
XX 26 ATGCAAGGCGCCACGCTCTGACCGCGCCACCTCTCTGCTGCGACCGCGCGCTG 85  
XX 21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGly 40  
XX 86 GCCGTGGCTCCCGGCGCTCGGTTCTGCTGACCGCCCGAGGATCATCAGGCCCGGAGA 145  
XX 41 AsnValThrIleGlyValGluLeuLeuGluHisCysProSerGlnValThrValIleAla 60  
XX 146 AATGTGACTATTGGGGTGGAGCTTCTGGAAACATGCGCTTCACAGGTGACTGAAGCG 205  
XX 61 GluLeuLeuLeuThrAlaSerAsnLeuThrValSerValLeuGluAlaGluGlyValPhe 80  
XX 206 GAGTGTCTCAAGACAGCATCAAACTCATCTCTCTGCTCGAAGACAGAGAGTCTTT 265  
XX 81 GluLysGlySerPheLeuThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGlu 100  
XX 266 GAAAGAGGCTTTTAAAGACCTTACTCTTCCATCATCTACTGAAACAGTGCAGATGAG 325  
XX 101 IleTyrGluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThr 120

XX 326 ATTTAAGCTACGCTGAACCGGACGTAACCGAGATGAGATTTTATCTCTAAATGATACC 385  
XX 121 ArgLeuSerPheGluThrLysArgIleSerValPheIleGlnThrAspLysAlaLeuTyr 140  
XX 386 CGCTTATCATTTAGACCAAGAGAAATTCGTCTTCAATTCAAACAGACAGCGCTTATAC 445  
XX 141 LysProLysGlnGluValLysPheArgIleValThrLeuPheSerAspPheLysProTyr 160  
XX 446 AAGCCAAAGCAAGAGAGGAGTTCCGATTTGATACCTCTTCAGATTTTAAAGCCTTAC 505  
XX 161 LysThrSerLeuLeuIleLeuIleLysAspProLysSerLeuLeuIleGlnIleProLeu 180  
XX 506 AAAACCTTTTAAACATTCATTAAAGACCCCAAAATTTGATTCACACAGTGGTTG 565  
XX 181 SerGlnIleSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIle 200  
XX 566 TCACAAACAAAGTATCTTGAGATCATTTCCAAAATTTTCAAGTATCTTCCATCCATATA 625  
XX 201 LeuGlyAspTyrPheSerIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGln 220  
XX 626 CTGTGTACTGCTCTATTCATCAAGTTCAAGTGAATGACACAGACATATCATCATTTACAG 685  
XX 221 ValSerGluTyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSer 240  
XX 686 GTTTCAGAAATGTATTAACCAAAATTTGAGTACCTTGGACACCATTAATATGTCTT 745  
XX 241 MetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProVal 260  
XX 746 ATGAATTTCAAGATTTTAAATGTATCATCATCAGCAAGATATCATATGGAAGCCAGTG 805  
XX 261 LysGlyAspValThrLeuThrPheLeuProLeuSerPheTyrGlyLysLysAsnIle 280  
XX 806 AAGAGAACGTAAACGCTTACATTTTAACTTTATCTTTTGGGGAAGAAAGAAATAATT 865  
XX 281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluGluMetLys 300  
XX 866 ACAAAAACATTTAAGATTAATGATGATCGCAACCTTCTTTAATGATGAAGAGATGAAA 925  
XX 301 AsnValMetAspSerSerAsnGlyLysSerGluTyrLeuAspLysSerProGlyPro 320  
XX 926 AATGTAAATGATTTCTTCAATGATGATCTTTCGAATACCTGATCTATCTTCCCTGAGCA 985  
XX 321 ValGluIleLeuThrThrValThrGluSerValThrGlyIleSerArgAsnValSerThr 340  
XX 986 GTAGAAATTTTAAACACAGAGACAGATCAGTATTCAGATTTTCAAGAAATGTAAGACT 1045  
XX 341 AsnValPhePheLysGlnHisAspTyrIleIleGluPheAspTyrThrThrValLeu 360  
XX 1046 AATGTGTCTTCAAGCAACATGATTAATCATTCATTTGATTAATTAATTAATCTACTGCTTG 1105  
XX 361 LysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsnGlnLeu 380  
XX 1106 AAGCATCTTCAACTTTCACAGCCACTGTGTAAGTAACTGTCTGATGCAACCAACTG 1165  
XX 381 ThrLeuGluGluArgArgAsnValIleIleThrValThrGlnArgAsnTyrThrGlu 400  
XX 1166 ACTCTTGAAGAAAGAAAGAAATTAATGATCAATACAGTGAACAGAGAAATCTACTGAG 1225  
XX 401 TyrTyrPheSerLysSerAsnSerGlyAsnGlnLysMetGluAlaValGlnLysIleAsnTyr 420  
XX 1226 TACTGAGCGGATCAATCAAGTGAATTAAGAAATGGAAGCTGTTTCAGAAATTAATTA 1285  
XX 421 ThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGluAspSerSerGlu 440  
XX 1286 ACTGTCCCAAGAGTGAACCTTTTAAGATTAATCCCAATCTCTGAGAGATCCAGTGAAG 1345  
XX 441 LeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSerLeuPhe 460  
XX 1346 CTACAGTGAAGGCTATTTCTTCTGTGATGAAGAAAGATGACAGTCACTATCTGTTT 1405  
XX 461 LysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLysValGly 480  
XX 1406 AAGTCTCTAGTAAGACATACATCCAAATTAAGAAAGATGATAATATTAAGGTGGGA 1465

QY 481 SerProheGluLeuValValSerGlyAsnLysArgLeuLeuSerTyrMetVal 500  
DB 1466 TGCCCTTTTGAAGTGGTGGTAGGCAACAAACGATTGAGAGATTGATATATGTA 1525  
QY 501 ValSerArgGlyGlnLeuValAlaValGlyGlnAsnSerTherMetSerLeuThr 520  
DB 1526 GATATCAGGGGACAGTGGTGGCTGTAGGAAAAACAAATTCACATGTTCTCTTAA 1585  
QY 521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyrTrillegLysAspGly 540  
DB 1586 CAGAAAAATTTCTGGACTCCAAAGCCTGTGATTTGTTATTTATTTAGAAATGATGG 1645  
QY 541 GlnIleLeSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560  
DB 1646 GAAATTATATAGATGTTCTAAAAAATTCCTGTTACGCTTTTAAAAAATAGATTAAG 1705  
QY 561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysAlaSerLeuArgIleSerVal 580  
DB 1706 CTATATTTGAGTAAAGTGAAGCTGAACCATCTGAGAAAAAGTCTCTTGAATCTCTGTG 1765  
QY 581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600  
DB 1766 AACACGCTGACTCCATAGTGGATTTGAGCTGTGACAAAGTGTGATCTGATGAAT 1825  
QY 601 AlaSerAsnAspIleTherMetGluAsnValValIleGlnLeuGluLeuTyrAsnThrGly 620  
DB 1826 GCCTCAATGATTTATGATGAAAAATGTGCTCAAGACTTGAACCTTTATTAACACAGA 1885  
QY 621 TyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTrpVal 640  
DB 1886 TATTTATTTAGCATGTTCAATGAAATCTTTGCAAGTCTTTCAGAAATGTGAGACTGGGTA 1945  
QY 641 LeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyrAspAsnAlaGluTyr 660  
DB 1946 TTGACAGATGCAAACTCAAGAAAGATTAATTTGAGTGGTTATACAAATCAGAAATAT 2005  
QY 661 AlaGluArgPheMetGluGlnLysGlnGlyIleValAspIleIleAspPheSerLeu 680  
DB 2006 GGTGAGAGGTTTATGAGGAAAAATGAGACACATATTGTGATATTCATGACTTTCTTGG 2065  
QY 681 GlySerSerProHisValArgLysHisPheProGluThrTrpIleTrpLeuAspThrAsn 700  
DB 2066 GGTAGACATCCACATGTCCGAAAGCATTTTCCAGACACTTGATTTGGCTAGACACAC 2125  
QY 701 MetGlyTyrArgIleTyrGlnGlnPheGluValThrValProAspSerIleThrSerTrp 720  
DB 2126 ATGGGTTNCAGATTTTACCAAGAAATTTGAAGTAACTGTAATCTTAATCACTTCTTGG 2185  
QY 721 ValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrProVal 740  
DB 2186 GTGGCTACTGGTTTGTGATCTCTGAGGACCTGGGCTTGGACTTAACACTCTCCAGTG 2245  
QY 741 GlnLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArg 760  
DB 2246 GAGCTCCAGACCTTCCACCATTTTTCATTTTTCATCTTCCACTCTGTATACAGA 2305  
QY 761 GlyGlnGlnPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrGluVal 780  
DB 2306 GGTGAGAAATTTGCTTTGAAAAATCAATATTAATTTGAAGATCCCACTGGGAT 2365  
QY 781 LysValIleIleGluLysSerAspLysPheAspIleLeuMetTherSerSerGluIleAsn 800  
DB 2366 AAGGATATCATTTGAGAAAAAGTACAAATTTGATTTCTAATGACTTCAATAGAAATTAAT 2425  
QY 801 AlaThrGlyHisGlnGlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPhe 820  
DB 2426 GCCACAGGCCACACAGACAGCCCTTGTGTTCCAGTGAAGATGGGCAACTGTCTTTT 2485  
QY 821 ProIleArgProThrHisLeuGlyGluIleProIleThrValIleThrAlaLeuSerProThr 840  
DB 2486 CCATCATGAGCAACATCTGGAGAAATTTCTATCATCAGTCAAGCTCTTACCCACT 2545

QY 841 AlaSerAspAlaValThrGlnMetIleLeuValLysAlaGlnGlyIleGluLysSerTyr 860  
DB 2546 GCTTCTGATGCTGTACCCACGATGATTTTATGTAAGGCTGAAGGATAGAAAAATCATAT 2605  
QY 861 SerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeu 880  
DB 2606 TCACATCATCTTAATTTAGACTGTGACTGACATAGGCTACAGAGTACCCCTAAAACTTGG 2665  
QY 881 SerPheSerPheProProAsnThrValThrGlySerGluAspValGlnIleThrAlaIle 900  
DB 2666 AGTTTCTATTTCTCTCAATACAGTACCTGGCAGTGAAGAAAGTTCAATCATCTGCAAT 2725  
QY 901 GlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGly 920  
DB 2726 GGAGATGTTCTTGGTCTTCCATCAATGGCTTAGCTCATTTGATTCGATGCTTATGGC 2785  
QY 921 CysGlyGlnGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThr 940  
DB 2786 TGTGTGAACAGAACATGATTAATTTGCTCCAAATATTTTACATTTTGGATTATCTGACT 2845  
QY 941 LysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPheMetArgGlnGly 960  
DB 2846 AAAAAGAAACAACTGACAGATTAATTTGAAAGAAAAAGCTTTTCAATTATAGAGCAAGCT 2905  
QY 961 TyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyr 980  
DB 2906 TACCAAGAGAAACTTCTTATCAAGGGAAGATGGCTTTTCAAGTCTTTTGGGAATAT 2965  
QY 981 AspProSerGlySerThrTrpLeuSerAlaPheValLeuArgCysPheLeuGluAlaAsp 1000  
DB 2966 GACCTTCTGGGAGACACTGTTGTGTCAGCTTTTGTTTAAAGTGTTCCTTGAAGCCGAT 3025  
QY 1001 ProTyrIleAspIleAspGluAsnValLeuHisArgThrTyrThrTrpLeuLysGlyHis 1020  
DB 3026 CTTTACATAGATTTGATCAGAAATGTGTTACACAGAAATACACTGTGCTTAAAGACAT 3085  
QY 1021 GlnLysSerAsnGlyGluPheTrpAspProGlyArgValIleHisSerGluLeuGlnGly 1040  
DB 3086 CAGAAATCCAAACGGTGAATTTTGGATCCAGAGAAAGTATTCATAGTACCTTCAAGT 3145  
QY 1041 GlyAsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArg 1060  
DB 3146 GGCATTAATAAGTCCGTAACACTTACAGCTATATTGTAACCTTCTCTCGGGAATATAGA 3205  
QY 1061 LysTyrGlnProAsnIleAspValGlnLysSerIleHisPheLeuGlnSerGluPheSer 1080  
DB 3206 AAGTATCAGCCTTAACATTTGATGTGCAGAGCTTATCCATTTTGGAGCTGGAATTCAGT 3265  
QY 1081 ArgGlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGly 1100  
DB 3266 AGAGGAATTTACAGCAATTAATTAATCTTAACCTTATTAATTTGATGTCAATCAAGTGGG 3325  
QY 1101 SerProLysAlaLysGluAlaLeuAsnMetLeuThrTrpArgAlaGlnGlnGluGly 1120  
DB 3326 AGTCTTAAGCAAGAAAGCTTTGAATATGCTGACTTGAAGACAGCAAGAAAGTGGC 3385  
QY 1121 MetGlnPheTrpValSerSerGluSerLysLeuSerAspSerTrpGlnProArgSerLeu 1140  
DB 3386 ATGCAATCTGGGTGTCTACAGAGTCCAAACCTTTCTGACTCTGGCACCCAGCTCCCTG 3445  
QY 1141 AspIleGluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGlu 1160  
DB 3446 GATATTTGAAGTTGACGCTTATGCACTGCTTCAACTTCTTAACATTTTCAAGCTTCTGAG 3505  
QY 1161 GlyIleProIleMetArgTrpLeuSerArgGlnArgAsnSerLeuGlyPheAlaSer 1180  
DB 3506 GGAATCCCAATTAAGAGTGGCTTAAGCAGCAAGAAATAGCTTGGGGTGGTTTGCACT 3565  
QY 1181 ThrGlnAspThrThrValAlaLeuLysValAlaLeuSerGluPheAlaAlaLeuMetAsnThr 1200  
DB 3566 ACTCAGATACACAGTGGCTTTTAAAGCTCTGTCTGAATTTGACAGCCCTTATAGATTAACA 3625  
QY 1201 GluArgThrAsnIleGlnValThrValThrGlyProSerSerProSerProLeuAlaVal 1220

```

Db      3626 GAAAGGACAAATATCCAGAGTACCGTGCAGCGGGGCTGACCAAGCTCTGCTGTG
Qy      1221 ValGlnProMetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeu
Db      3686 GTACAGCCCAACGGCGATTAAATATTCGCCAAATGGTTTGGATTGCTATTGTCAGCTC
Qy      1241 AsnValValTyrAsnValIleAlaSerGlySerSerArgArgArgSerIleGlnAsn
Db      3746 AATGTTGATATATATGTGAAGGCTTCTGCTTCTAGAAAGACGAAGTATCATCAAAAT
Qy      1261 GlnGlnAlaPheAspLeuAspValAlaValIleGlnLeuAsnIleAspAspLeuAsnHisVal
Db      3806 CAAGAAAGCTTTGATTAGATGTGCTGTAAAGAAATAAAGATGATCTCAATCATGTG
Qy      1281 AspleuAsnValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGln
Db      3866 GATTGGAATGTGTGTACAAAGCTTTCCGGCCCGGTAGAGATGGCATGCTTTATATGGA
Qy      1301 ValAsnLeuLeuSerGlyPheMetValProSerGlnAlaIleSerLeuSerGluThrVal
Db      3926 GTTAACCTATTAAAGTGGCTTTATGGTGCTTCAGAAAGCAATTTCTCTGAGGAGACAGTG
Qy      1321 LysIleValGluTyrAspHisGlyLysLeuAsnLeuTyrLeuAspSerValAsnGluThr
Db      3986 AAGAAAGGGAATGTGATCATGAAACTCAACCTATTAGATTCTGTAAATGAAACC
Qy      1341 GlnPheCysValAsnIleProAlaValArgAsnIleValSerAsnThrGlnAspAla
Db      4046 CAGTTTGTGTAAATATTCCTGCTGTGAAACTTTAAAGTTTCAAAATCCCAAGATGCT
Qy      1361 SerValSerIleValAspTyrTyrGlnProArgArgGlnAlaValArgSerTyrAsnSer
Db      4106 TCAGTGTCATATAGTGATTAATAGAGCCAAAGAGACAGCGGTGAAGATTAACAAGCTC
Qy      1381 GluValIleLeuSerSerCysAspLeuCysSerAspValGlnGlyCysArgProCysGln
Db      4166 GAAAGTGAAGCTGTCTCTGTCAGACCTTTCAGATGCCAGGCTGCCCTTGTGTAG
Qy      1401 AspGlyIleSerGlySerHisHisIleSerSerValIlePheIlePheCysPheLeu
Db      4226 GATGAGGCTTCAAGGCTCCATCATCATCTTCAGCATTTTATTTCTGTTCAGCTT
Qy      1421 LeuTyrPheMetGluLeuTyrLeu
Db      4286 CTGTACTTATGGAAGCTTGGCTG
RESULT 2
ID      AAL49815 standard; cDNA; 4335 BP.
XX      AAL49815;
XX      AAL49815;
XX      04-DEC-2002 (first entry)
XX      Human platelet alloantigen Gova coding sequence.
XX      Human: platelet alloantigen: Gova; Govb: single nucleotide polymorphism;
XX      SNr: diallelic; bleeding disorder; post-transfusion purpura;
XX      Km post-transfusion platelet refractoriness; haemostatic; vaccine;
XX      neonatal alloimmune thrombocytopenia; gene; ss.
OS      Homo sapiens.
XX      XX
FH      Key Location/Qualifiers
FT      CDS 1..4335
FT      FT /*tag= a
FT      FT /product= "Gova"
FT      FT /partial
FT      FT /note= "no stop codon"
FT      FT replace(2108,C)
FT      FT /*tag= b
XX      XX

```

```

PN      W0200270738-A2.
XX      12-SEP-2002.
XX      07-MAR-2002; 2002W0-CA00291.
XX      07-MAR-2001; 2001US-273941P.
XX      (SCHU//) SCHUH A.
XX      PI Schuh A, Ouwehand W;
XX      WPI; 2002-713460/77.
XX      DR P-PSDB; AAO19372.
XX      PT New isolated oligonucleotide binding to a region of CD109 nucleic acid
XX      PT having a single nucleotide polymorphism that distinguishes a Gova
XX      PT and/or Govb allele, useful for treating blood disorders e.g. alloimmune
XX      PT thrombocytopenia.
XX      PS Claim 6; Page 23-29; 69pp; English.
XX      CC The present invention relates to a sequence capable of binding
XX      CC specifically to a CD109 nucleic acid which has a single nucleotide
XX      CC polymorphism that distinguishes the Gova and Govb alleles. Detection of
XX      CC the Gov genotype is useful for detecting whether the subject has or is at
XX      CC risk of a blood disease, disorder or abnormal physical state, such as
XX      CC bleeding, or increased risk of bleeding, due to alloimmune destruction of
XX      CC blood platelets, e.g., post-transfusion purpura, post-transfusion
XX      CC platelet refractoriness or neonatal alloimmune thrombocytopenia. The
XX      CC nucleic acid and polypeptide are useful for Gov genotyping or phenotyping
XX      CC individuals. The present sequence is the Gova coding sequence.
SQ      Sequence 4335 BP; 1296 A; 837 C; 922 G; 1280 T; 0 other:
Alignment Scores:
Pred. No.: 0 Length: 4335
Score: 7326.50 Matches: 1427
Percent Similarity: 98.82% Conservative: 1
Best Local Similarity: 98.75% Mismatches: 0
Query Match: 99.71% Indels: 17
DB: 24 Gaps: 1
US-10-020-095-4 (1-1428) x AAL49815 (1-4335)
Qy      1 MetGlnGlyProProLeuLeuThrAlaHisIleLeuCysValCysThrAlaIleLeu
Db      1 ATGAGGAGCCACCGCTCTGACCGCGCCACCTCTCTGCGTGCACCGCGCGCTG
Qy      21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGlyGly
Db      61 GCCGCGCTCCGGGCGCTGCTTCTGTGTAGACCCGACGGATCATCAGCGCGAGGA
Qy      41 AsnValThrIleGlyValGluLeuLeuGlnHisCysProSerGlnValThrValIleVal
Db      121 AATGTGACTATTGGGCTGAGGCTTCTGGAACCTGCCCTTCAACAGGTGATGGAAGCGG
Qy      61 GluLeuLeuYsrThrIleSerAsnLeuThrValSerValLeuGlnIleGlnIleValPhe
Db      181 GAGCTGTCAAGACAGGATTAACCTCACTGTCTGTCTGTGAAGCAGAGAGTCTTT
Qy      81 GluIleGlySerPheIleYsrThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGlu
Db      241 GAABAAGGCTCTTTAAAGACACTTACTCTTCCATCATCACTGTAAGACGTGAGATGAG
Qy      101 IleTyrGluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThr
Db      301 ATTATAGAGCTACGTGAACCGGACCGATCAAGAGATGAGATTTATTTCTTAATAGTACC
Qy      121 ArgLeuSerPheGlnThrIleYsrArgIleSerValPheIleGlnThrAspIleValLeuTyr
Db      361 CGCTATCATTTGAGACCAAGAGAAATATGTCTTCAATTCACAGACCAAGGCTTATAC

```

QY 141 LysProLysGlnGluValLysPheArgIleValThrLeuPheSerAspPheLysProTyr 160  
DB 421 AAGCCAAAGCAAGAGAGAGTTCCGATGTGTAACCTCTTCAGATTTTAAGCCTTAC 480  
QY 161 LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGlnTyrPheLeu 180  
DB 481 AAAACCTCTTAAACATCTCAATTAAAGACCCCAAAATTCAAATTTGATTCAAAGCTGTTG 540  
QY 181 SerGlnInSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIle 200  
DB 541 TCACAAACAAGATCATCTGAGATCATTTCCAAAACCTTTCAGCTATCTTCCCATCCAAAT 600  
QY 201 LeuGlyAspTyrSerIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGln 220  
DB 601 CTGGGAGACGTGCTCTATTCAGGTTCAGGTAAAGACACAGCATATATCATCATTTTCAG 660  
QY 221 ValSerGluTyrAlaLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSer 240  
DB 661 GTTTCAGAAATATATATATACCAAAATTTGAAGTACCTTGACACCATTTATATGTCTT 720  
QY 241 MetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProVal 260  
DB 721 ATGAATCTTACAGATTTAAATGATACATACCGCAAAAGATATCATATGGAAAGCCAGTG 780  
QY 261 LysGlyAspValThrLeuThrPheLeuProLeuSerPheTyrGlyLysLysAsnIle 280  
DB 781 AAAGAGACGTACGCTTACATTTTATCCTTTATCCTTTGGGAAAGAAAGAAATATTT 840  
QY 281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluGluMetLys 300  
DB 841 ACAAAAACATTTAAGATTAATGATCTGCMAACTTCTCTTTATATATATGAAAGATGAAA 900  
QY 301 AsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSerSerProGlyPro 320  
DB 901 AATGTATATGATCTTCAATGACATTTCTGAAATACCTGAGATCTATCTTCCCTGGACCA 960  
QY 321 ValGluIleLeuThrThrValThrGluSerValThrGlyIleSerArgAsnValSerThr 340  
DB 961 GTGGAATTTTAAACACAGTACAGATCAGTATACAGATTTTCAAGAAATGTAAACACT 1020  
QY 341 AsnValPhePheLysGlnHisAspTyrIleIleGluPheAspTyrThrThrValLeu 360  
DB 1021 AAGTGTCTTCAAGACACATGATTCATCATTCATGAGTTTTCATTATATCTACTGCTTG 1080  
QY 361 LysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsnGlnLeu 380  
DB 1081 AAGCCATCTCTCAACTTCACAGCCACTGGAAGGTAACTGTCTGATGTCAACCAACTG 1140  
QY 381 ThrLeuGluGluArgArgAsnAsnValIleThrValThrGlnArgAsnTyrThrGlu 400  
DB 1141 ACTCTTGAAGAAAGAAATATATGTATGATACATACATGACACAGAAACATATCTGAG 1200  
QY 401 TyrTyrSerGlySerAsnSerGlyAsnGlnLysMetGluAlaValGlnLysIleAsnTyr 420  
DB 1201 TACTGAGCGGATCTAAAGTGAATTCAGAAATTCAGAACTGTTCAGAAATTAATTAAT 1260  
QY 421 ThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGlnAspSerSerGlu 440  
DB 1261 ACTGTCCCAAGAGAGAACTTTTAAGATTGAATTCCTCAATCTCGAGAGATTCACATGAG 1320  
QY 441 LeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSerLeuPhe 460  
DB 1321 CTACATTTGAAGCCCTATTTCTTGGTAGTAAAGTAGACATGGCAGTTCAATCTGCTTT 1380  
QY 461 LysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLysValGly 480  
DB 1381 AAGTCTCTTGAAGATCATCATCCAACTAAACAAAGATGAAATATATTAAGGTGGAG 1440  
QY 481 SerProPheGluLeuValLysSerGlyAsnLysArgLeuLysLeuLysSerTyrMetVal 500  
DB 1441 TCCTCTTTTGAAGTTGGTGTAGTGGCAAAACGATGAAAGAGTTTAAGCTTATTAAGTGA 1500  
QY 501 ValSerArgGlyGlnLeuValAlaValGlyLysGlnAsnSerThrMetPheSerLeuThr 520

DB 1501 GATATCAGGGGACAGTTGGCTGCTAGAGAAAACAAATTCACAATGTCTCTTAAACA 1560  
QY 521 ProGluAsnSerTyrThrProLysAlaCysValIleValTyrTyrIleGluAspAspGly 540  
DB 1561 CCAGAAATTTCTTGACTTCAAAAGCTGTGTAAATGTGTATATATGAAGATGATGGG 1620  
QY 541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560  
DB 1621 GAAATTTATAGATGATGTTCTTAAATTCCTGTTCAGCTTGTTTTAAATATAGATTAAG 1680  
QY 561 LeuTyrTyrSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580  
DB 1681 CTATATTTGAGTAAAGTGAAGCTGAACCATCTGAAGAAAGTCTCTTGTAGATCTGTG 1740  
QY 581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600  
DB 1741 ACACAGCTGACTCCATATGTGGATGTGAGCTGTGACAAAAGTGTGAATCTGATGAAT 1800  
QY 601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeuGluLeuTyrAsnThrGly 620  
DB 1801 GCCTCTAATGATATTTACATGAGAAATGTGTCCATGAGTTGGAATCTTATTAACACAGA 1860  
QY 621 TyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTyrVal 640  
DB 1861 TATTTATTTAGCATGTTCATGAATCTTTTGCAGCTTTCAAGAAATGTGACTGTGGATA 1920  
QY 641 LeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyrAspAsnAlaGluTyr 660  
DB 1921 TTGACAGATGCAAACTCCAGAAAGATTAATATGATGGTTATGACATATCAGAAATAT 1980  
QY 661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680  
DB 1981 GCTGAGAGTTTATGAGAGAAATGAGACATATTTGATATTCATATCATCTTTCTTGG 2040  
QY 681 GlySerSerProHisValArgLysHisPheProGluThrTyrIleTyrPheAspThrAsn 700  
DB 2041 GGTAGCAGTCCACATGTCCGAAGACTTTTCCAGAGACTTGTGATTTGGCTAGACACAAAC 2100  
QY 701 MetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIleThrSerTyr 720  
DB 2101 ATGGGTATACAGATTTTACCAAAATTTGAAGTACGTACATCTGATCTATCTCTGG 2160  
QY 721 ValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrThrProVal 740  
DB 2161 GTGGCTACTGTTTGTGATCTCTGAGACCTGTGGCTTGGACTTACCAACTCTCCAGTG 2220  
QY 741 GluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArg 760  
DB 2221 GAGCTCAAGCCTTCCACATTTTTCATTTTGTGAATCTTCCCTACTCTGTATACAA 2280  
QY 761 GlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrGluVal 780  
DB 2281 GGTGAAGATTTGCTTGGAAATTAATATATTAATTAATTAATTAATTAATTAATTAAT 2340  
QY 781 LysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsn 800  
DB 2341 AAGGTATATGAGAAAGTGAACAATTTGATATCTTATATACCTCAATATGAATAAAT 2400  
QY 801 AlaThrGlyHisGlnGlnThrLeuLeuValProSerGluAspGlyValAlaThrValLeuPhe 820  
DB 2401 GCCACAGGCCACACAGACACCTTCTGTCTCCCATGAGAGAGGGGCAACTGTCTTTT 2460  
QY 821 ProIleArgProThrHisLeuGlyGluIleProIleThrValThrAlaLeuSerProThr 840  
DB 2461 CCCATCAGGCCAACATCTGGAGAAATTTCTATACACAGTCAAGCTCTTCAACCCACT 2520  
QY 841 AlaSerAspAlaValThrGluMetIleLeuValLysValGluGlyIleGluLysSerTyr 860  
DB 2521 GCTTGTAGTGTGTACCCAGATGATTTTATGTAAGGCTGAAGAGATGAATAAATCATAT 2580  
QY 861 SerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeu 880



[illegible]

OY		1224	MetalavalanilleserAlaengilyPheglyPhealalleCysginLeuAsnValVal	1243
Db		3721	ATGGAGATTAAATATTTCCGCAAAAGTGTTCGATTCTCATTTCTACGCTCAAGTGTGA	3780
OY		1244	TyrAsnVallysalaserGlySerSerArgRgaRgRserIleGlnAsnGluAla	1263
Db		3781	TATAATGAAAAGGCTCTCGGGTCTTCTAGAAGACGAAGATTAATCAAAATCAAGAAC	3840
OY		1264	PheAspleuAepValalaVallyGluAsnLysAspAspleuAsnlsvalAspleuAsn	1283
Db		3841	TTTGATTTAGATGTTGCCTGTAAGAAAAATPAAGATGATCTCATTCATGTGCATTTGAT	3900
OY		1284	ValCystrHsrerPheSerGlyProGlyARgSerGlyMetAlaleuMetGluValAsnLeu	1303
Db		3901	GTCGTACAACCTTTTCGGGCCCGGTAGAGATGGCATGGCTCTTATGAAAGTTAACCTA	3960
OY		1304	LeuSerGlyPheMetValProSerGluAlalleSerLeuSerGluThrVallyVal	1323
Db		3961	TTAAGTGCTTTAATAGTGCTTCAGAAACAATTTCTCTGAGCGACAGTGAAGAAAGTG	4020
OY		1324	GluTyraAspHisgllyLysLeuAsnLeuTyrlasuAsperValaengluthrGlnPheCys	1343
Db		4021	GAATTAATGATCATGGAAGAACTCAACTCTATTAGATTCTGTAATGAAGAACCGTTTTGT	4080
OY		1344	ValAsnilleProAlaValArgAsnPhelyValSerAnthrGlnAspAlaserValSer	1363
Db		4081	GTTAATATTTCTCGCTGTGAGAAACTTTAAAGTTCCAATATCCCAAGATGCTTCAGTCTCC	4140
OY		1364	IleValAspyrryrrGluProARgArgAlaValargserTyranSerGluValLys	1383
Db		4141	ATAGGAGTAACTVATGAGCCAAAGAGACAGCGGTGAGAAGTAACTCTGAAGTAGAAC	4200
OY		1384	LeuSerSerCysAspLeuCysSerAspValGlnglyCysArgProCysGluAspGlyAla	1403
Db		4201	CTGTCCTCTGTGACCCTTGAGAGTAGTCCAGSGCTGCCCTCTTGTGAGGATGGAGCT	4260
OY		1404	SerglySerHisHisHisSerSerValillePheHellePhecysPheylsleuleUtyrPhe	1423
Db		4261	TCAGGCTCCCATCATCACTCTTCAGTCATTTTTTTTCTGTGTTCAAGCTTCTGTACTTT	4320
OY		1424	MetGluLeuThrleu 1428	
Db		4321	ATGGAACCTTGGCTTG 4335	
<b>RESULT 3</b>				
ID	ABO79964			
XX	ABO79964 standard; cDNA; 4761 BP.			
XX	ABO79964;			
DT	23-DEC-2002 (first entry)			
DE	Human CD109 K1 protein encoding cDNA.			
KW	CD109; alpha2 macroglobulin; thioester; cerebroprotective; cardiant; immunosuppressive; haemostatic; anticoagulant; thrombolytic; human; cardiovascular; vasotropic; gene therapy; CD109 K1; gene; ss.			
OS	Homo sapiens.			
FH	Key	Location/Qualifiers		
FT	CDS	113..4450		
FT		/tag= a		
FT		/product= "CD109 K1"		
PN	WO200270696-A2.			
PD	12-SEP-2002.			
PF	07-MAR-2002; 2002WO-CA00292.			
PR	07-MAR-2001; 2001US-273814P.			

XX (SCHU/) SCHU A.  
PA (SUTH/) SUTHERLAND R D.  
XX PI Schuh A. Sutherland RD;  
XX WPI; 2002-713450/77.  
DR P-PSDB; ABB82165.

XX New CD109 nucleic acids and polypeptides, useful in gene therapy,  
PT particularly for treating strokes, myocardial infarctions, thrombosis,  
PT thrombocytopenia, autoimmune diseases, or organ or bone marrow  
PT transplantation -

XX Claim 1; Fig 1a; 156pp; English.

XX The invention relates to isolated nucleic acid molecules encoding CD109  
CC polypeptides. These nucleic acid molecules include the human CDNA  
CC sequences comprising CD109 K1, CD109 K1-H7, CD109 K15 or their variants.  
CC CD109 is a novel member of the alpha2 macroglobulin (alpha2M)/C3, C4, C5  
CC family of cholesteryl-containing proteins. The CD109 polypeptides can be  
CC expressed by standard recombinant methodology. The CD109 nucleic acid,  
CC CD109 polypeptide or its mimetic, or the anti-CD109 antibody is useful  
CC for treating or detecting a disease or disorder, e.g. conditions  
CC associated with endothelial activation, platelet activation, activation  
CC of the coagulation or fibrinolytic systems, activation of T lymphocytes  
CC and of the complement system, quantitative or qualitative abnormalities  
CC of platelet function, increased or impaired platelet aggregation and/or  
CC activation, increased or impaired activation of the coagulation and/or  
CC fibrinolytic systems, or impaired or increased immune activation. These  
CC are also useful for treating cardiovascular disorders, stroke, myocardial  
CC infarction, thrombosis, embolism, peripheral vascular disease,  
CC thrombocytopenia, chomobocytchemia, autoimmune diseases, organ  
CC transplantation, or bone marrow transplantation. The CD109 nucleic acids  
CC or their antisense nucleotide sequence are useful in gene therapy for  
CC treating these conditions. The present sequence represents the human  
CC CD109 K1 cDNA sequence.

XX Sequence 4761 BP; 1420 A; 912 C; 1016 G; 1413 T; 0 other;

XX Alignment Scores:

Pred. No.: 0 Length: 4761  
Score: 7326.50 Matches: 1427  
Percent Similarity: 98.82% Conservative: 1  
Best Local Similarity: 98.75% Mismatches: 0  
Query Match: 99.71% Indels: 17  
DB: 24 Gaps: 1

US-10-020-095-4 (1-1428) x ABO79964 (1-4761)

QY 1 MetGInGlyProProleuLeuThrAlaAlaIleuLeuValCysThrAlaAlaLeu 20  
DB 113 ATGCGAGGCGCACCGCTCTGACCGCGCCACCTCTCTGCGTGCACCGCGGCGTG 172  
QY 21 AlaValAlaProGlyProArGpPheLeuValThnAlaProGlyIleIleArGpProGlyGly 40  
DB 173 GCGGCGGCTCCGGGCGCTCGGTTTCGTGACAGCCCGCAGGATCATCGCCGAGGA 232  
QY 41 AenValThrIleGlyValAlaLeuLeuGluHsCysProSerGlnValThrVallySala 60  
DB 233 AATGAGACTATTTGGGGGTGAGGCTTCGGAACACTGCTTCAACAGTGACTGGAAGCGG 292  
QY 61 GluLeuLeuValThralaSerAenLeuThralaSerValLeuGluAlaGluGlyValPhe 80  
DB 293 GAGCTGCTCAAGACAGCATCAAACTCACTGCTCTCTGCTGAGACAGAGAGAGCTTT 352  
QY 81 GluLySgLySerPheLySerThrLeuThrLeuProSerLeuProLeuAenSerAlaaspGlu 100  
DB 353 GAAAAAGGCTCTTTTAAGACACTTACTCTTCATCATCACTACTCTGAAAGTGAGTGA 412  
QY 101 IleTyrGluLeuAlaGlyAlaThrGlyAlaGlyThrGlnaspGluIleLeuPheSerAenSerThr 120  
DB 413 AATTATGAGCTACGTGTAAACCGACGTAACCCAGATGAGATTTATTCTTAATAGTACC 472

QY 121 ArgLeuSerPheGluThrThryAlaArgIleSerValPheIleGlnThrAspLysAlaLeuTyr 140  
DB 473 CGCTTATCATTTGAGACCAAGAGAAATATCTGCTTCATTCAAAAGACAGCCCTTATAC 532  
QY 141 LysProLyGInGluValAllyPheArGllEValThrLeuPheSerAspPheLySProTyr 160  
DB 533 AAGCCAAAGCAAGAAAGTGAAGTTGCGACATTTGTTACACTCTCTCGAATTTTAAAGCTTAC 592  
QY 161 LysThrSerLeuAenIleleuIleLysAaspProLySerAspLeuIleGlnGlnProLeu 180  
DB 593 AAAACCTCTTAAACATCTCATTAAGAGACCCCAATCAAAATTTAATCAACAGAGTGTTG 652  
QY 181 SerGInGInSerAspLeuGlyValIleSerLyThrPheGlnLeuSerSerHisProIle 200  
DB 653 TCACAAACAAGGACTCTTGAGTCACTTTCCAAAACCTTTACGCTATCTTCCATCCATTA 712  
QY 201 LeuGlyAapTTPSerIleGlnValAlaAenaspGlnThrTyrTyrGlnSerPheGln 220  
DB 713 CTGGGTGACTGGTCTATTCAAGTTCAAGTGAATGACACAGACATATTATCAATCATTTGAG 772  
QY 221 ValSerGluTyrValLeuProLyPheGlnValThrLeuGlnThrProLeuTyrCysSer 240  
DB 773 GTTCAAGAAATGATATTAACCAAAATTTAAGAGACTTTGACAGACACATTAATATTGTTCT 832  
QY 241 MetAenSerLySHisLeuAenGlyThrIleThralaLySThrTyrGlyLysProVal 260  
DB 833 ATGAATTTCAAGCATTTTAAATGCTACATCAACGGAAGTATACATATGGAAGGCAAGT 892  
QY 261 LysGlyAapValThrLeuThrPheLeuProLeuSerPheTyrGlyLysAlaAenIle 280  
DB 893 AAAGAGAGCGTAAACCTTACATTTTACCTTATCTTTGGGGAAGAAAGAAATATTT 952  
QY 281 ThrLySThrPheLySHisLeuAenGlySerAlaAenPheSerPheAenaspGluIleMetLys 300  
DB 953 ACAAAAACATTTAAATAAATGATGATCTCAAACTCTCTTTTAAATGAAGAAAGATGAAA 1012  
QY 301 AenValMetAapSerSerAsnGlyLeuSerGluTyrLeuAaspLeuSerSerProGlyPro 320  
DB 1013 AATGTAAATGATATCTTCAAAATGACTTCTGTAATCTCGAGTATCTTCCCTGAGACA 1072  
QY 321 ValGluIleLeuThrThrValThrGluSerValThrGlyIleSerArgAenValSerThr 340  
DB 1073 GTAGAAATTTTAAACACAGTGACAGAAATCAGTTACAGTTATTTCAAGAAATGAAGCACT 1132  
QY 341 AenValPhePheLyGlnHisAspTyrIleIleGluPhePheAspTyrThrThrValLeu 360  
DB 1133 AATGTGTTCTTCAACCAACATGATTAATCAATGATTTGATTTGATTATTAATCTCTTG 1192  
QY 361 LysProSerLeuAenPheThralaSerValThrVallySValThrArgAlaaspGlyAenGlnLeu 380  
DB 1193 AAGCCATCTCTCACTTCAACAGCCACTGTGAAGGAAAGTACCTGCGTGAAGGCAACCACTG 1252  
QY 381 ThrLeuGluGluArgArgAenAenValValIleThrValThnGlnArgAenTyrThrGlu 400  
DB 1253 ACTCTTGAAGAAAGAAATATATGTAAGTACATTAACAGTACACAGAAATATATCTGAG 1312  
QY 401 TyrTTPSerGlySerAenSerGlyAenGlnLySmetGluAlaValGlnLySleAsnTyr 420  
DB 1313 TACTGAGACGGAATCAACAGTGAATCAAGAAAGGAAGCGTTCAAGAAATTAATTAAT 1372  
QY 421 ThrValProGlnSerGlyThrPheLySleGluPheProIleLeuGluAaspSerSerGlu 440  
DB 1373 ACTGTCCTCCAAAGTGAAGCTTTTAAGTTGATTTGATTTCCAAATCTTGAAGATTTCCAGTGA 1432  
QY 441 LeuGlnLeuLySAlaTyrPheLeuGlySerLySserSerMetAlaValHisSerLeuPhe 460  
DB 1433 CTACAGTTGAAGGCTTATTTCTTGTTAGTAAAGTATACAGAGGAGTTCATATGCTGTTT 1492  
QY 461 LysSerProSerLySThrTyrIleGlnLeuLySThrArgaspGluAenIleLySValGly 480  
DB 1493 AAGTCTCTTAGTAAAGACATACATCACTAAACCAAGATGAAATTAATTAAGGTGGA 1552

Qy 481 SerProhegIuLeuValIaSerGIyAsnLysaRgLeuLysGIuLeuSerTyRMeCVal 500  
Db 1553 TCGCCTTTTGGATTGGTGGTAGTGGCAACAACGATTGAAGAGTTAAGCTATATGGTA 1612  
Qy 501 ValSerArgGIuGIuLeuValIaValGIyLysGIuAsnSerThRMePheSerLeuThr 520  
Db 1613 GTATCCAGGGAGCAGTTGGTGGCTGTAGAGAAACAAATTCACAAAGTTCTCTTTACA 1672  
Qy 521 ProGIuAsnSerThRProLysAlaCySerValIleValTyRTrIleGIuAspArgGIy 540  
Db 1673 CCAAGAAATTTCTGGACCTCCAAAAGCTGTAAATGTGTATTTATTTAGAGATGAGGG 1732  
Qy 541 GluIleIleSerAspValIleuLysIleProValGIuLeuValPheLysAsnLysIleLys 560  
Db 1733 GAAATTTATTAAGTATGTTCTTAAAAATTCCTGTTCAGCTGTTTAAAAATTAAGATAAG 1792  
Qy 561 LeuTyRTrpSerLysValIleValAGIuProSerGIuLysValSerLeuArgIleSerVal 580  
Db 1793 CTATATTTGGAGTAAAGTAAAGCTGAACCATCTGAGAAAGCTCTTTAGGATCTGTG 1852  
Qy 581 ThrGIuProAspSerIleValGIyIleValIaValAspLysSerValAsnLeuMeCAsn 600  
Db 1853 ACAAGCCTGACTCATAGTGGTGGATTTAGCTGTTCACAAAAGTGTCAATCTGATGAT 1912  
Qy 601 AlaSerAsnAspIleThRMeCGLuAsnValIaIleGIuLeuGIuLeuTyRAsnThRGIy 620  
Db 1913 GCCTCTATATGATATTAACAATGGAATAATGGTCCATGAGTTGGAACCTTATACACAGA 1972  
Qy 621 TyRTrIleuGIuMeCPhMeCAsnSerPheAlaValPheGIuGIuCyRGIyLeuTrpVal 640  
Db 1973 TATATTTAGGACGATTCATGAATTCCTTTTCAGCTTTTCAGGAATGGAGCTCGGGTA 2032  
Qy 641 LeuThRAspAlaAsnLeuThRAspTyRTrIleAspGIyValTyRAspAsnAlaGIuTyR 660  
Db 2033 TTGACAGATGCAACCTTCACAGAGATTTATTTAGTGGTGTATTCAGACATGCGAATAT 2092  
Qy 661 AlaGIuArgPheMeCGLuGIuAsnGIuGIyIleValIaAspIleHisAspPheSerLeu 680  
Db 2093 GCTGAGGCTTTATGGAGGAAATGAAGACATATTTGATATTCATGACTTTTCTTTG 2152  
Qy 681 GlySerSerProHisValArgLysHisPheProGIuThRTrpIleThRLeuAspThRAsn 700  
Db 2153 GGTAGCAGTCCACATGTCGGAAGAGATTTTCAGAGACTTGGATTTGGCTAGACACAAAC 2212  
Qy 701 MetGIyTyRArgTrIeTyRGIuGIuPheGIuValThRValProAspSerIleThRSerTrp 720  
Db 2213 ATGGGTTACAGGATTTACCAAGAAATTTGAAGTAATGACTGATCTATCATCTTCTGG 2272  
Qy 721 ValAlaThRGIyPheValIleSerGIuAspLeuGIuLeuGIyLeuThRThRProVal 740  
Db 2273 GTGGCTACTGGTTTGTGATCTCTGAGAGACCTGGGCTTTGGACTTACAACTACTCCAGTG 2332  
Qy 741 GluLeuGIuAlaPheGIuProPhePheIlePheLeuAsnLeuProTyRSerValIleArg 760  
Db 2333 GAGCTCCCAAGCCTTCCAAACCATTTTTCATTTTTCATTTTTCCTCTCTCTGTATCAGA 2392  
Qy 761 GlyGIuGIuPheAlaLeuGIuIleThRTrIlePheAsnTyRLeuLysAspAlaThRGIuVal 780  
Db 2393 GGTAAAGAAATTTGCTTGGAAATTAACAATTAATTCATTAATTTGAAGAGCCACGAGGTT 2452  
Qy 781 LysValIleIleGIuLysSerAspLysPheAspIleLeuMeCThRSerSerGIuIleAsn 800  
Db 2453 AAGGTAAATTCATGGAAAGTGAACAATTTGATATTTCAATGACTTCAATGAATTAAT 2512  
Qy 801 AlaThRGIyHisGIuGIuThRLeuLeuValProSerGIuAspGIyAlaThRValLeuPhe 820  
Db 2513 GCCACAGGCCACACAGCAACCTTCTGTGTTCCAGTGGAGATGGGGCAACTGTTCTTTT 2572  
Qy 821 ProIleArgProThRHisLeuGIuGIuLeuProIleThRValThRAlaLeuSerProThr 840  
Db 2573 CCCATCCAGGCCAACATCTGGAGAAATTTCTTATCACAGTCAAGCTTTTTCACCACT 2632  
Qy 841 AlaSerAspAlaValThRGIuMeCILEuValLysAlaGIuGIyIleGIuLysSerTyR 860

Db 2633 GCTTCGATGCTGCTCACCCAGATGATTTAGTAAGGCTGCAAGAAATAGAAAATCATAT 2692  
Qy 861 SerGIuSerIleLeuLeuAspLeuThRAspAsnArgLeuGIuInSerThRLeuLysThRLeu 880  
Db 2693 TCACATTCATCTTATTAAGCTTGACTGACAAATAGGCTACAGAGTACCTGAAAATCTTGG 2752  
Qy 881 SerPheSerPheProAsnThRValThRGIySerGIuArgValGIuIleThRAlaIle 900  
Db 2753 AGTTTTCATTTCTCTTAATACGTGACTGGCAGTGAAGAGTTCAGATCATGCAATT 2812  
Qy 901 GlyAspValLeuGIuProSerIleAsnGIyLeuAlaSerLeuIleArgMeCProTyRGIy 920  
Db 2813 GGAAGATTTCTTGGTCTTCATCAATAGGCTTGACCTCATTTGATTTGGAGTTCATGCGC 2872  
Qy 921 CysGIyGIuGIuAsnMeCILEuAsnPheAlaProAsnIleTyRTrIleLeuAspTyRLeuThr 940  
Db 2873 TGTGGTAAACAGAACATGATTAATTTTGCTCCAAATTTTACATTTTGGATTTATCTGACT 2932  
Qy 941 LysLysLysGIuInLeuThRAspAsnLeuLysGIuLysAlaLeuSerPheMeCArgGIy 960  
Db 2933 AAAAAGAAACACTGACAGATTAATTTGAAGAAAAGCTCTTCATTATGAGGCAAGGT 2992  
Qy 961 TyRGIuArgGIuLeuLeuTyRGIuArgGIuAspGIySerPheSerAlaPheGIyAsnTyR 980  
Db 2993 TACCAGAGAGAACTTCTATCAGAGGGAAAGATGGCTTTTCAGTCTTTTGGAAATTAAT 3052  
Qy 981 AspProSerGIySerThRTrpLeuSerAlaPheValLeuArgLysPheLeuGIuAlaAsp 1000  
Db 3053 GACCTTTCTGGAGCACTGGTGTGTGAGCTTTGTTTAAAGATTTCTTTGAAGCCGAT 3112  
Qy 1001 ProTyRTrIleAspIleAspGIuAsnValLeuHisArgThRTrIleThRTrpLeuLysGIyHis 1020  
Db 3113 CTTTACATATGATTTGATCGAATGTGTTCACAGAAACATACCTTGGCTTAAGAGCAT 3172  
Qy 1021 GluLysSerAsnGIyLysPheThRAspProGIyArgValIleHisSerGIuLeuGIuGIy 1040  
Db 3173 CAGAAATCCAAACGCTGATTTGGGATCCAGAGAGTGAATTCATGATGAGCTTCAAGGT 3232  
Qy 1041 GlyAsnLysSerProValThRLeuThRAlaTyRTrIleValHisSerLeuLeuGIyTyRArg 1060  
Db 3233 GGCATTAATAAGTCCAGTAACCTTACAGCTTAATTTGTAATCTTCTCCGGGAAATAGA 3292  
Qy 1061 LysTyRGIuProAsnIleAspValGIuGIuSerTrIleHisPheLeuGIuSerGIuPheSer 1080  
Db 3293 AAGTATCAGCCTTAACATGATGTGCAAGACTATCATTTTGTGAGTGTGAATTCAGT 3352  
Qy 1081 ArgGIyTrIleSerAspAsnTyRThRLeuAlaLeuIleThRTrAlaLeuSerSerValGIy 1100  
Db 3353 AGAGGAATTTCAACAATTAATTAATCTTACGCTTATTAATCATTAATGATCATCAGTGGGG 3412  
Qy 1101 SerProLysAlaLysGIuAlaLeuAsnMeCLeuThRTrpArgAlaGIuGIuGIuGIy 1120  
Db 3413 AGTCTTAACCGAAGAGAGCTTTGAATATGCTACTTGAGACACAGACAAAGAGGGGC 3472  
Qy 1121 MetGIuPheThRValSerSerGIuSerLysLeuSerAspSerTrpGIuProArgSerLeu 1140  
Db 3473 ATGCAATTTCTGGTGTTCATAGAGTCCAAACTTTTGACTCCCTGGCAGCCACCTCCCTG 3532  
Qy 1141 AspIleGIuValAlaAlaTyRAlaLeuLeuSerHisPheLeuGIuPheGIuThRSerGIu 1160  
Db 3533 GATATTAAGATTTGAGCTGATGCACTGCTCACACTTCTTAATTTAGACTTCTGAG 3592  
Qy 1161 GlyTrIleProIleMetArgTrpLeuSerArgGIuArgAsnSerLeuGIyGIyPheAlaSer 1180  
Db 3593 GGAATCCCAATTTATGAGTGGCTTAACAGGCAAGAAATAGCTTGGTGGTGTTCATCT 3652  
Qy 1181 ThRGIuAspThRThRValAlaLeuLysAlaLeuSerGIuPheAlaAlaLeuMeCAsnThR 1200  
Db 3653 ACTCAGGATACCACTGTGGCTTTTAAAGCTCTGTCTGAATTTTCAGCCCTTAAGAAATACA 3712  
Qy 1201 GluArgThRAsnIleGIuValThRValThRGIyProSerSerProSerPro----- 1217

```
Db 3713 GAAAGGACAAATATCCAGTGAACCGTGAAGGCGCTTACGTCACCAAGTCTCTGTAAGTTT 3772
Qy 1218 -----LeuAlaValAlaGlnPro 1223
Db 3773 CTGATTGACACACAAACCGCTTACTCTTACAGACAGACAGCTGCTGTGGTGAACGCCA 3832
Qy 1224 MetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValAla 1243
Db 3833 ATGGCAGTTAAATATATCCCAATAGTGTGATTTGATTTGTCAGCTCAATGTTCTGA 3892
Qy 1244 TyrAsnValIleAlaSerGlySerSerArgArgArgArgSerIleGlnAsnGlnGlnAla 1263
Db 3893 TATTAATGTAAGGCTTCTGGGCTTCTTACAAAGACGAAATCTATCCAAATCAAGAAAGCC 3952
Qy 1264 PheAspLeuAspValAlaValIleGlyIleAsnIleAspAspLeuAsnIleValAspLeuAsn 1283
Db 3953 TTTGATTGATGATGTTGCTGTAAAGAAATTAAGATGATCTCATGATGATGATTTGAT 4012
Qy 1284 ValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGlnValAsnLeu 1303
Db 4013 GTGTGTACAAAGCTTTTCGGGCGCGGTAGAGAGTGCATGCTTATGGAAGTTAACTTA 4072
Qy 1304 LeuSerGlyPheMetValProSerGlyAlaIleSerLeuSerGlnThrValIleGlyVal 1323
Db 4073 TTTAAGTGGCTTATAGTGGCTTCAAGAGCAATTTCTTACGCGAAGCAGTGAAGAAAGTG 4132
Qy 1334 GluTyrAspPheIleGlyIleLeuAsnLeuTyrLeuAspSerValAsnGlnThrGlnPheCys 1343
Db 4133 GAATATGATGATGAAAGAACTCAACCTCTATTTAGATCTGTAAATGAACCCAGTTTGT 4192
Qy 1344 ValAsnIleProAlaValArgAsnPheIleSerAsnThrGlnAspAlaSerValSer 1363
Db 4193 GTTAATATTTCCGCTGTCAGAACTTTAAAGTTTCAATATACCAGATGCTTCAAGTGTCC 4252
Qy 1364 IleValAspTyrTyrGlnProArgArgGlnAlaValArgSerTyrAsnSerGlnValIle 1383
Db 4253 ATAGTGTGATTAATGAGCCAAAGAGACAGCGGTGAGAAATTAACAATCTGAAAGTGAAG 4312
Qy 1384 LeuSerSerCysAspLeuCysSerSerAspValGlnGlyCysArgProCysGlnAspGlyAla 1403
Db 4313 CTGTCTCTCTGACCTTTGCGATGATGATGCGAGGCTGCGCTTGTGATGATGAGT 4372
Qy 1404 SerGlySerHisHisSerSerValIlePheIlePheCysPheIleLeuTyrPhe 1423
Db 4373 TCAGGCTCCCATCATCATCTTCAAGTATTTTATTTCTGTTTCAAGCTTCTGTACTTT 4432
Qy 1424 MetGlnLeuTyrLeu 1428
Db 4433 ATGGAACCTTGGCTG 4447
RESULT 4
AAD49440
ID AAD49440 standard, DNA; 5882 BP.
XX AAD49440;
XX
XX 24-MAR-2003 (first entry)
XX
XX Human blood cell surface antigen, CD109 encoding DNA #2.
DE Human blood cell surface antigen, CD109 encoding DNA #2.
XX
XX Human; GPI-anchored TGF-beta1 binding protein; r150 protein; cancer;
KW glycosylphosphatidyl inositol; transforming growth factor-beta1;
XX therapy; blood cell surface antigen; CD109; ds.
OS Homo sapiens.
XX
XX WO200285942-A2.
XX
XX 31-OCT-2002.
XX
XX 24-APR-2002; 2002WO-CA00560.
XX
XX 24-APR-2001; 2001US-285713P.
```

```
PR 14-FEB-2002; 2002US-356163P.
XX
XX (UWMC-) UNIV MCGILL.
XX
XX Philip A, Tam B;
XX
XX WPI; 2003-093100/08.
XX
XX Novel transforming growth factor (TGF)-beta 1 binding reagent which
PT comprises r150 protein which acts as accessory receptor of TGF-beta,
PT useful for negatively modulating TGF-beta activity, and thus for
PT treating cancer
XX
XX Disclosure; Fig 17; 127bp; English.
XX
XX The invention relates to novel transforming growth factor (TGF)-beta1
CC binding reagent which comprises a glycosylphosphatidyl inositol (GPI)-
CC anchored TGF-beta1 binding protein referred to as r150 which acts as
CC an accessory receptor of TGF-beta. The invention is used for negatively
CC modulating TGF-beta activity, and thus for treating conditions
CC characterised by overproduction of TGF-beta, such as cancer. Antisense
CC molecules of the invention are used for increasing TGF-beta availability
CC and increase graft success. The present sequence is human blood cell
CC surface antigen, CD109 encoding DNA. CD109 is a variant of r150 sequence.
CC Note: This sequence is stated to encode human CD109 protein shown in
CC figure 17 of the specification, but this does not appear to be the case.
SQ Sequence 5882 BP; 1696 A; 1173 C; 1241 G; 1772 T; 0 other;
Alignment Scores:
Pred. No.: 0 Length: 5882
Score: 7326.50 Matches: 1427
Percent Similarity: 98.82% Conservative: 1
Best Local Similarity: 98.75% Mismatches: 0
Query Match: 99.71% Indels: 17
DB: Gaps: 1
US-10-020-095-4 (1-1428) x AAD49440 (1-5882)
Qy 1 MetGlnGlyProProLeuLeuThrAlaHisIleLeuCysValCysThrAlaAlaLeu 20
Db 113 ATGCAAGGCGCCACCGCTCTGACCGCGCCACCTCTGTGTGTGACCGCCCGCTG 172
Qy 21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGly 40
Db 173 GCCGTGGCTCCGAGCTCTGCTTGTGTGACGCCCGCAGGATATCAAGCCCGAGGA 232
Qy 41 AsnValThrIleGlyValGlnLeuLeuGlnHisCysProSerGlnValThrValIleAla 60
Db 233 AATGTGACTAATTGGGGTGGAGCTTCTGGAACACCTGCCCTTCAACAGTACTGTGAAGGG 292
Qy 61 GluLeuLeuTyrThrAlaSerAsnLeuThrValSerValLeuGlnAlaGlnGlyValPhe 80
Db 293 GAGCTGCTCAAGACAGCATCAAACTCACTGCTCTGCGAGAGCAGAAAGAGTCTTT 352
Qy 81 GluTyrGlySerPheIleThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGlu 100
Db 353 GAAAAGGCTCTTTAAAGACACTTACTCTTCATCACTCACTCAAGACGTCAACATGAG 412
Qy 101 IleTyrGlnLeuArgValThrGlyArgThrGlnAspGlnIleLeuPheSerAsnSerThr 120
Db 413 ATTTATGAGCTACGCTGAACCGAGATGCCAGATGAGATTTTCTCTAAATAGTACC 472
Qy 121 ArgLeuSerPheGlnThrLysArgIleSerValPheIleGlnThrAspIleAlaLeuTyr 140
Db 473 CGCTTATCATTTGAACCAAGAGATATCTGCTTCATTCAAACAGACAAAGCCCTTAAAC 532
Qy 141 LysProIleGlnGlnValIlePheArgIleValThrLeuPheSerAspPheIleProTyr 160
Db 533 AAGCCAAAGCAAGAGTAAAGTTTGCATGTGTAACCTCTCTCAATTTAAAGCTTAAAC 592
Qy 161 LysThrSerLeuAsnIleLeuIleLysAspProIleSerAsnLeuIleGlnGlnThrLeu 180
```

Db 593 AAAACCTCTTAAACATCTCATTAAAGACCCCAATCAATTGATCCAACTGGTGG 652  
Qy SerGlnGlnSerAspLeuGlyValIleSerIleThrPheGlnLeuSerSerHisProIle 200  
Db 653 TCACAAACAAAGTGAATCTTGGAGTCATTTCCAAACCTTTTCAGCTATCTTCCCATCCAAATA 712  
Qy LeuGlyAspTyrSerIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGln 220  
Db 713 CTGGTGACTGGTCTATTCAGATTCAGATGCAATGACGACATATTCATTCATTCATTCAG 772  
Qy ValSerGlyTyrValLeuProIlePheGlnValIleThrLeuGlnThrProLeuTyrCysSer 240  
Db 773 GTTTCAGAAATATGATTAACCAAAATTTGAACTGCTTGGACAGACCAATTAATTTGTTCT 832  
Qy MetAsnSerIleValIleLeuAsnGlyThrIleThrAlaIleValIleValIleValIleValIle 260  
Db 833 ATGAAATTCATGCAATTTAAATGTAACATTCAGGCAAAATGTAATGTAATGTAATGTAATG 892  
Qy LysGlyAspValIleThrLeuThrPheLeuProLeuSerPheTyrGlyValIleValIle 280  
Db 893 AAAGAGACGTAAGCTTACATTTTACCTTTATCTTTTGGGAAAGAAAGAAATATTT 952  
Qy ThrIleThrPheIleValIleValIleValIleValIleValIleValIleValIleValIle 300  
Db 953 ACAAACCAATTAAGATTAATGATTCGCAAACTTCTTTTAATGATGAAGATGATAA 1012  
Qy AsnValMetAspSerSerAsnGlyLeuSerGlyTyrIleLeuAspLeuSerSerProGlyPro 320  
Db 1013 AATGTAATGATCTTCAATGACCTTTCTGAATACCTGATCTGATCTTCTTCTTCTGAGACA 1072  
Qy ValGlyIleLeuThrThrValIleThrGlnSerValIleThrGlyIleSerAspAsnValSerThr 340  
Db 1073 GTACAAATTTTAACACAGTACAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 1132  
Qy AsnValIlePheIleValIleValIleValIleValIleValIleValIleValIleValIle 360  
Db 1133 AATGATGCTTCAAGCAACATGATTAATCAATGATGATGATGATGATGATGATGATGATG 1192  
Qy LysProSerLeuAsnPheThrAlaThrValIleValIleValIleValIleValIleValIle 380  
Db 1193 AAGCATCTCTCAATCTTCAAGCACTGTAAGGATTCGTCGTCGTCGTCGTCGTCGTCGTCG 1252  
Qy ThrLeuGlnGlyAspAsnAsnValIleThrValIleThrGlnAspAsnValSerThrGlu 400  
Db 1253 ACTCTTGAAGAAAGAAATTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 1312  
Qy TyrTrpSerGlySerAsnSerGlyAsnGlnIleMetGluValIleGlnIleValIleValIle 420  
Db 1313 TACTGAGCCGATCTTAACAGTGAATCAAGAAATGGAAGCTGTCGTAAGAAATTAATG 1372  
Qy ThrValIleProGlnSerGlyThrPheIleIleGluPheProIleLeuGlnAspSerSerGlu 440  
Db 1373 ACTGCCCCCAAGGAACTTTTAAGATTAATTTCCAAATCTCGAGGATTCAGATTCAGAT 1432  
Qy LeuGlnLeuValAlaIlePheLeuGlySerIleSerSerMetAlaValIleSerIlePhe 460  
Db 1433 CTACAGTTGAAGGCTATTTCTTGGTAGTAAGTAGTACATGAGATTCAGATTCAGATTCAG 1492  
Qy LysSerProSerIleThrTyrIleGlnLeuValIleThrAspGluAsnIleValIleValIle 480  
Db 1493 AAGCTCTCTACTAAGACATACATCCAACTAAACCAAGATGAAATTAATTAAGTGGCA 1552  
Qy SerProPheGlnLeuValIleSerGlyAsnValIleValIleValIleValIleValIleVal 500  
Db 1553 TCGGCTTTTGAAGTGGTGTAGTGAACAAACCAATGGAAGGATTAAGCTTAATGATTA 1612  
Qy ValSerArgGlyGlnLeuValAlaValIleValIleValIleValIleValIleValIleVal 520  
Db 1613 GTATCCAGGAGGACGTTGGTGTGAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTA 1672  
Qy ProGlnAsnSerTrpThrProIleValIleValIleValIleValIleValIleValIleVal 540  
Db 1673 CCAAAATATCTTGGACTCCAAAGCCGTGTAATTTGATTAATTTGAAAGATGATGGG 1732

Qy GlnIleIleSerAspValIleLeuValIleProValIleGlnLeuValIlePheIleValIleVal 560  
Db 1733 GAATATTAATGATGATGTTCTTAATAAATTTCCGTTCCAGCTGTTCTTTTAAATAATTAAGATTAAG 1792  
Qy LeuTyrTrpSerIleValIleValIleValIleValIleValIleValIleValIleValIle 580  
Db 1793 CTATATTTGAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTA 1852  
Qy ThrGlnProAspSerIleValIleValIleValIleValIleValIleValIleValIleVal 600  
Db 1853 ACACAGCTTACCTCAGTATGAGATTTGAGCTGTTGACAAAGTGTGAATGTAATGTAATG 1912  
Qy AlAspAsnAspIleThrMetGluAsnValIleValIleGlnLeuGlnLeuTyrAsnThrGly 620  
Db 1913 GCCTTAATGATATTTCAATGAAAGTAATGTCATGATGATGATGATGATGATGATGATGAT 1972  
Qy TyrTyrLeuGlyMetPheMetAsnSerPheAlaValIlePheGlnGlnIleValIleValIle 640  
Db 1973 TATTTATTTAGCATGTTTCATGAATTTCTTTTGGAGTCTTTCAGGAATGTCAGATCTGGGTA 2032  
Qy LeuThrAspAlaLeuLeuThrIleValIleValIleValIleValIleValIleValIleVal 660  
Db 2033 TTGACAGATCCAAACCTCAGAAAGATTAATGATGATGATGATGATGATGATGATGATGAT 2092  
Qy AlaGlyArgPheMetGluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 680  
Db 2093 GCTGAGAGGTTTATGAGAGAAATGAAAGACATTTGTAATGATTAATGATTAATGATTAATG 2152  
Qy GlySerSerProIleValIleValIleValIleValIleValIleValIleValIleValIle 700  
Db 2153 GTTAGCAGTCCACATGTCGCAAAAGCATTTTCCAGAGCTGGAATTTGGCTAGACACCAAC 2212  
Qy MetGlyTyrArgIleTyrGlnGlnIlePheGlnValIleThrValProAspSerIleThrSerTrp 720  
Db 2213 ATGGGTACAGGATTTACAAAGATTTGAAGTAACGTGATCTGATCTGATCTGATCTGATCT 2272  
Qy ValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrThrProVal 740  
Db 2273 GTGGCTACTGTTTGTGATCTGAGAGACTGGCTTGTGATTAACATTAATCCAGCTCCAGTG 2332  
Qy GlnLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArg 760  
Db 2333 GAGCTCCAAACCTTCCAAACCATTTTCAATTTTTCATTTTTCATTTTTCATTTTTCATTT 2392  
Qy GlyGlnGlnPheAlaLeuGlnIleThrIlePheAsnTyrLeuIleValIleValIleValIle 780  
Db 2393 GGTGAAGAAATTTCTTTGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2452  
Qy IleValIleIleGlnIleValIleValIleValIleValIleValIleValIleValIleVal 800  
Db 2453 AAGGTATCTATTTGAAGAAATGACCAATTTGATTTCTATTTGATTTCTCAATGAATTAAT 2512  
Qy AlaThrGlyIleGlnGlnIleThrLeuValIleProSerGluAspGlyAlaThrValIlePhe 820  
Db 2513 GCACAGGACCAAGGACGAGCCCTTCTGCTGCTCCAGTGAAGATGGGCAACGTTCTTTT 2572  
Qy ProIleArgProThrIleLeuGlyGlnIleProIleThrValIleThrAlaLeuSerProThr 840  
Db 2573 CCATCAGGCAACATCTGGAAGAAATTCATATCACTGACAGTCTTTTCAACCCACT 2632  
Qy AlAspAspAlaValIleThrGlnMetIleLeuValIleValIleValIleValIleValIleVal 860  
Db 2633 GCTTCTATGCTGACCCAGATGATTTTGTAAAGCTGAAGAAATGAAGAAATTAATTAAT 2692  
Qy SerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuValIleThrLeu 880  
Db 2693 TCACATCCATCTTATTAAGATGACATGACATTAAGGATCAAGATCAAGATCAAGATCAAG 2752  
Qy SerPheSerPheProProAsnThrValIleThrGlySerGlnArgValIleGlnIleThrAlaIle 900  
Db 2753 AGTTTCATATTTCTCTTAATTAAGTACGTCGAGTGAAGAGATTCAGATCACTGCAATT 2812



DR WPI: 2002-713450/77.  
 DR P-PSDB; ABB82167.  
 XX New CD109 nucleic acids and polypeptides, useful in gene therapy,  
 PT particularly for treating strokes, myocardial infarctions, thrombosis,  
 PT thrombocytopenia, autoimmune diseases, or organ or bone marrow  
 PT transplantation -  
 XX  
 PS Claim 1; Fig 2a; 156pp; English.  
 XX  
 CC The invention relates to isolated nucleic acid molecules encoding CD109  
 CC polypeptides. These nucleic acid molecules include the human cDNA  
 CC sequences comprising CD109 K1, CD109 K1-H7, CD109 K15 or their variants.  
 CC CD109 is a novel member of the alpha2 macroglobulin (alpha2M)/C3, C4, C5  
 CC family of thioester-containing proteins. The CD109 polypeptides can be  
 CC expressed by standard recombinant methodology. The CD109 nucleic acid,  
 CC CD109 polypeptide or its mimetic, or the anti-CD109 antibody is useful  
 CC for treating or detecting a disease or disorder, e.g. conditions  
 CC associated with endothelial activation, platelet activation, activation  
 CC of the coagulation or fibrinolytic systems, activation of T lymphocytes  
 CC and of the complement system, quantitative or qualitative abnormalities  
 CC of platelet function, increased or impaired platelet aggregation and/or  
 CC activation, increased or impaired activation of the coagulation and/or  
 CC fibrinolytic systems, or impaired or increased immune activation. These  
 CC are also useful for treating cardiovascular disorders, stroke, myocardial  
 CC infarction, thrombosis, embolism, peripheral vascular disease,  
 CC thrombocytopenia, thrombocythemia, autoimmune diseases, organ  
 CC transplantation, or bone marrow transplantation. The CD109 nucleic acids  
 CC or their antisense nucleotide sequence are useful in gene therapy for  
 CC treating these conditions. The present sequence represents the human  
 CC CD109 K1-H7 cDNA sequence.  
 XX  
 SQ Sequence 5895 BP; 1709 A; 1173 C; 1241 G; 1772 T; 0 other;

## Alignment Scores:

Pred. No.: 0 Length: 5895  
 Score: 7326.50 Matches: 1427  
 Percent Similarity: 98.82% Conservative: 1  
 Best Local Similarity: 98.75% Mismatches: 0  
 Query Match: 99.71% Indels: 17  
 DB: 24 Gaps: 1

US-10-020-095-4 (1-1428) x ABQ79966 (1-5895)

QY 1 MetGlnGlyProProLeuLeuThrAlaIahIshLeuLeuCyValCysThrAlaIaleu 20  
 DB 113 ATGAGGGCCCAACGCTCTGACGCGCCCAACCTCTCTGCTGGTGACGCGCGCTG 172  
 QY 21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGlyGly 40  
 DB 173 GCCCGGCTCCCGGCGCTCGGTTTCTGGTGACAGCCCAAGGATCATCAGGCCGAGGA 232  
 QY 41 AsnValThrIleGlyValGluLeuLeuGluHisCysProSerGlnValThrValIleAla 60  
 DB 233 AATGACTATTTGGGGTGGAGCTTCTGGAAACATGCGCTTCAAGTACGTGAGCG 292  
 QY 61 GluLeuLeuLeuThrAlaSerAsnLeuThrValSerValLeuGluAlaGluGlyValPhe 80  
 DB 293 GAGCTGCTCAAGACGACATCAAACTCACTGCTCTGCTGGAAGCAAGAGAGTCTT 352  
 QY 81 GluIysGlySerPheLeuThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGlu 100  
 DB 353 GAAAGAGGCTCTTTAAACACCTTACTTCTTCATCATCACTCTGAAACAGTGCAGATGAG 412  
 QY 101 IleTyrgluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThr 120  
 DB 413 ATTATAGACTACGCTGAACCGAGTACCCAGAGTGAATTTATTTCTTAATATGATACC 472  
 QY 121 ArgLeuSerPheGluThrIleArgIleSerValPheIleGlnThrAspValAlaLeuTyr 140  
 DB 473 CGCTTATCATTTGAGACCAAGAAATATCTGCTTCAATCAAAAGCAAGAGCCCTTAATAC 532  
 QY 141 LysProIysGlnGluValIleValPheArgIleValThrLeuPheSerAspPheLeuProTyr 160

DB 533 AAGCCAAGACAAAGAGGATTGCGATTGTTACACTCTTCTCAGATTTTAAACCTTAC 592  
 QY 161 LysThrSerLeuAsnIleLeuIleValAspProIysSerAsnLeuIleGlnThrLeu 180  
 DB 593 AAAACCTCTTAAACATTTCAATTAAGAACCCCAAAATCAAAATTTGATTCACACAGTGGTTG 652  
 QY 181 SerGlnIserAspLeuGlyValIleSerIysThrPheGlnLeuSerSerHisProIle 200  
 DB 653 TCACAAACAAAGTATCTTGGAGTATTTCCAAAACCTTTAGCATCTTCCATCCACATA 712  
 QY 201 LeuGlyAspTrpSerIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGln 220  
 DB 713 CTGGTACATCGTCTATTCAAGTTCAAGTAATGACACGACATTTATCATCATTTTCAG 772  
 QY 221 ValSerGluTyrValLeuProIysPheGluValThrLeuGlnThrProLeuTyrCysSer 240  
 DB 773 GTTTCAGAAATATGATATCAAAAATTTGAAGTGCTTGGACAGACACATATATGATTTCT 832  
 QY 241 MetAsnSerIysHisLeuAsnGlyThrIleThrAlaIysTyrThrTyrGlyLysProVal 260  
 DB 833 ATGAATTTCAACATTTAAATGATGACATCAGCCAAAGTATATGGAAGCCAGTG 892  
 QY 261 LysGlyAspValThrLeuThrPheLeuProLeuSerPheTrpGlyLysIleValAsnIle 280  
 DB 893 AAAGAGACGTAAACGCTTACATTTTACCTTTATCCTTTGGGGAAGAAAGAAATATTT 952  
 QY 281 ThrIysThrPheIysIleAsnGlySerAlaAsnPheSerPheAsnAspGluGluMetLys 300  
 DB 953 ACAAACCAATTTAAGATTAATGATGATGCAAACTCTCTTTATATGATGAAGATATAA 1012  
 QY 301 AsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSerSerProGlyPro 320  
 DB 1013 AATGTATGATTTCTTCAATATGACCTTCTGAATCACTGGATCATCTTCCCGAGACA 1072  
 QY 321 ValGluIleLeuThrThrValThrGluSerValThrGlyIleSerArgAsnValSerThr 340  
 DB 1073 GTAGAAATTTTAAACACAGACGACAGATCAAGTTACAGGTATTTCAAAAATGTAAAGACT 1132  
 QY 341 AsnValPhePheIysGlnHisAspTyrIleIleGluPhePheAspTyrThrThrValLeu 360  
 DB 1133 AATGTGTTCTTCAAGCAACATGATTCATCATGATTTTGTATATACACTAGTCTTG 1192  
 QY 361 LysProSerLeuAsnPheThrAlaThrValIleValThrArgAlaAspGlyAsnGlnLeu 380  
 DB 1193 AAGCATCTCTCAACTTCAACAGCACCTGTAAGSTACTGCTGANTGCAACCACTG 1252  
 QY 381 ThrLeuGluIuArgArgAsnAsnValIleIleThrValThrGlnArgAsnTyrThrGlu 400  
 DB 1253 ACTCTGAAGAAAGAAATTAATGTAGTCATTAACAGTACACAGAGAAACTATACAGAG 1312  
 QY 401 TyrTrpSerGlySerAsnSerGlyAsnGlnIleMetCysIleValIleGlnIleAsnTyr 420  
 DB 1313 TACTGAGCGGATCTTAACAGTGAATCAAGAAATGGAACCTGTCGAAAAATTAATATAT 1372  
 QY 421 ThrValProGlnSerGlyThrPheIysIleGluPheProIleLeuGluAspSerSerGlu 440  
 DB 1373 ACTGTCCCCCAAACTGAACTTTAAGATTGAATTCCTCAATCTCGAGAGATTCACAGTAC 1432  
 QY 441 LeuGlnLeuIysAlaTyrPheLeuGlySerIysSerSerMetAlaValHisSerLeuPhe 460  
 DB 1433 CTACAGTTGAAGGCTATTTCTTGTGTAATAAGTACATGAGCATTAAGTCTGTTT 1492  
 QY 461 LysSerProSerIysThrTyrTyrIleGlnLeuLeuThrArgAspGluAsnIleValGly 480  
 DB 1493 AAGTCTCTTATGATACATCACTCACTAAACCAAGATGAAAAATATAAAGGTGGGA 1552  
 QY 481 SerProPheGluLeuValIleSerGlyAsnIysArgLeuIysGluLeuSerTyrMetVal 500  
 DB 1553 TCGCTTTTGAAGTGGTGGTTAGTGGCAAAAGATTTGAAGAGTTAAGCTATATGATA 1612  
 QY 501 ValSerArgIysGlnLeuValAlaValGlyLysGlnAsnSerThrMetPheSerLeuThr 520



```
Db      1613 GATTCACGGGACACGTTGGCTGTAGAGAAAACAAATTCACAAATGTTCTCTTAACA 1672
Qy      521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGluAspArgIy 540
Db      1673 CCAGAAATATCTTGGATCCCAAAAGCCTGTGTATTTGTATATTTAGAAATGATGG 1732
Qy      541 GluIleIleSerAspValLeuLysIleProValGluLeuValPheLysAsnLysIleLys 560
Db      1733 GAAATATATAGCATGCTTCTAAATAATCTCTGTCACCTGTTTAAAAATAGATAAAG 1792
Qy      561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580
Db      1793 CTATATGGAGTAAAGTGAAGCTGAACATCTGAAGAAAGCTCTCTTGAGATCTCTGG 1852
Qy      581 ThrGluProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600
Db      1853 AACAGCCTGATCCCATAGTTGGATTTGAGCTGTGACAAAAGTGTGATCTGATGAAT 1912
Qy      601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeuGluLeuTyrAsnThrGly 620
Db      1913 GCCTTAATGATATTCAAATGGAATAATGTGTCATGAGTTGGAACTTTATACACAGA 1972
Qy      621 TyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGluGluCysGlyLeuTrpVal 640
Db      1973 TATATTTAGGCAATGTTCAATGAAATCTTTTGCACTTTTCAGAAATGTGACTGGGTA 2032
Qy      641 LeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyrAspAsnAlaGluTyr 660
Db      2033 TTGACAGATGCACAACTCCACGAAGATTTATATTGATGAGTGTATTATACAAATGCAGAA 2092
Qy      661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680
Db      2093 GCTGAGAGGTTTATGAGAGAAATGAAAGACATATTGTAGATATTCATGACTTTCTTGG 2152
Qy      681 GlySerSerProHisValArgLysHisPheProGluThrTrpIleTrpLeuAspThrAsn 700
Db      2153 GGATGAGATCCACATGTCCGAAAGCAATTTTCCAGACACTGTGATTTGGCTAGACACMAAC 2212
Qy      701 MetGlyTyrArgIleTyrGluGluPheGluValThrValProAspSerIleThrSerTrp 720
Db      2213 ATGGGTTACAGATTTTACCAAGAAATTTGAAGTACGTACCTGATTTCTATCTCTTGTGG 2272
Qy      721 ValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrThrProVal 740
Db      2273 GGGGCTACTGGTTTGTGATCTCTGAGACCTGGGCTTGGACTAACCACTACTCCAGAG 2332
Qy      741 GluLeuGluAlaPheGluProPhePheIlePheLeuAsnLeuProTyrSerValIleArg 760
Db      2333 GAGCTCCAAAGCCTTCCAAACATTTTTCATTTTGTGAATCTTCCCTACTCTGTATCAGA 2392
Qy      761 GlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrGluVal 780
Db      2393 GGGTAAAGAAATTTGCTTGGAAATTAATAATTAATTAATTTGAATAATCCACTGGGTT 2452
Qy      781 LysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsn 800
Db      2453 AAGGTATCATTTGAGAAAGTAGCAAAATTTGATATTCTATATGACTTCAATGAAATTAAT 2512
Qy      801 AlaThrGlyHisGluGluThrLeuLeuValProSerGluAspGlyValaThrValLeuPhe 820
Db      2513 GCCACAGGCGCACGACAGACCTTCTGTGTCCAGAGTAGAGATGGGCAACTGTTCTTTT 2572
Qy      821 ProIleArgProThrHisLeuGlyGluIleProIleThrValThrAlaLeuSerProThr 840
Db      2573 CCATATAGGCACACATCTGGAGAAATTCATATCAGTACACACACTCTTTCACCCACT 2632
Qy      841 AlaSerAspAlaValThrGluMetIleLeuValIleGluGlyIleGluLysSerTyr 860
Db      2633 GCTTCGATGCTGTCAACCAAGATTTTATGTAAGCTGGAAGAAATTAATAATATATAT 2692
Qy      861 SerGluSerIleLeuLeuAspLeuThrAspAsnArgLeuGluSerThrLeuLysThrLeu 880
Db      2693 TCACATATCATCTTATTAGACTTGAACATATAGGCTACAGAGTACCTCGAAAACTTTG 2752

Qy      881 SerPheSerPheProAsnThrValThrGlySerGluArgValGluIleThrAlaIle 900
Db      2753 AGTTTCTATTTCCCTCAATATACATGACTGGCACTGAAAGATTCACATCTCTCAATT 2812
Qy      901 GlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGly 920
Db      2813 GGAGATGTTCTTGTGCTTCCATCAATGAGCTTAGCCTCATGTGATTCGAGATGCTTATGGC 2872
Qy      921 CysGlyGluGluAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThr 940
Db      2873 TGTGTGAAACAGAAATATTAATTTGCTCCAAATATTTATATTTGATATATCTGACT 2932
Qy      941 LysLysLysGluLeuThrAspAsnLeuLysGluLysAlaLeuSerPheMetArgGluGly 960
Db      2933 AAAAAGAAACAACATGACATGATATTTGAAAAGAAAAGCTCTTTCATTTATATGAGCAAGT 2992
Qy      961 TyrGluArgGluLeuLeuTyrGluArgGluAspGlySerPheSerAlaPheGlyAsnTyr 980
Db      2993 TACACAGAGAACTTCTCATACAGAGAGATGGCTCTTTCAGTGGCTTTTGGAAATTA 3052
Qy      981 AspProSerGlySerThrTrpLeuSerAlaPheValLeuArgCysPheLeuGluAlaAsp 1000
Db      3053 GACCTTCTGGAGACATTTGTTGTTCACGCTTTTGTTTAAGATGTTTCTTGAAACCGAT 3112
Qy      1001 ProTyrIleAspIleAspGluAsnValLeuHisArgThrTyrThrTrpLeuLysGlyHis 1020
Db      3113 CCTTACATAGATATTGATCAGAAATGTGTTACACAGAACATACACTTGGCTTTAAGACAT 3172
Qy      1021 GlnLysSerAsnGlyGluPheTrpAspProGlyArgValIleHisSerGluLeuGluGly 1040
Db      3173 CAGAAATCCACAGGGAATTTTGGATCCAGAGAGATGATCATATGAGCTTCAAGT 3232
Qy      1041 GlyAsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArg 1060
Db      3233 GGCAATTAAGCCAGTAACCTTACACCATATATTTGAATCTTCTCTCGGATATGGA 3292
Qy      1061 LysTyrGluProAsnIleAspValGluGluSerIleHisPheLeuGluSerGluPheSer 1080
Db      3293 AAGTATCAGCTTACATTTGATGTGCAAGATCTATCCATTTTGTGAGTCTGAATTCAGT 3352
Qy      1081 ArgGlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGly 1100
Db      3353 AGAGAAATTTTCAACAATTAATATCTTACCTTAACTTATATGATTTGCATCAGTGGGG 3412
Qy      1101 SerProLysAlaLysGluAlaLeuAsnMetLeuThrTrpArgAlaGluGluGlyGly 1120
Db      3413 AGTCTTAAAGCGAAGAGCTTTGAAATGTGCTGACTTGAAGAGCAGAACAAAGAGTGGC 3472
Qy      1121 MetGluPheTrpValSerSerGluSerLysLeuSerAspSerTrpGluProArgSerLeu 1140
Db      3473 ATGCATTTCTGGGTCTATCAGAGTCCAAACTTCTGTGCTCTGCGACACGCTCCCTG 3532
Qy      1141 AspIleGluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGluPheGluThrSerGlu 1160
Db      3533 GATATTTGAAGTTGACGCTTATGACCTGTCTCAGACTTCTTACATTTTCAGACTTCTAG 3592
Qy      1161 GlyIleProIleMetArgTrpLeuSerArgGluArgAsnSerLeuGlyPheAlaSer 1180
Db      3593 GGAAATCCCAATATATGAGGTGGCTTAAGCAGCAAGAAATACTGGGCTGTTTGGCATCT 3652
Qy      1181 ThrGluAspThrThrValAlaLeuLysValAlaLeuSerGluPheAlaAlaLeuMetAsnThr 1200
Db      3653 ACTCAGATACACATGCTGTGCTTTAAAGCTCTGTGTGAATTTGACGCCCTTAATGAATCA 3712
Qy      1201 GluArgThrAsnIleGluValaThrValThrGlyProSerSerProSerPro----- 1217
Db      3713 GAAAGGACAAATATATCAAGTACCGTGAACGGGCTTACTCACCAAGTCTCTTAAGTTT 3772
Qy      1218 -----LeuAlaValaGluPro 1223
Db      3773 CTGATTTGACACACAAACCGTTACTCTTGCAGACAGACAGACTTGTGTGTGTCACGCCA 3832
```

QY 1224 MetAlaValAsnIleSerAlaAsnGlyPheAlaIleCysGlnLeuAsnVal 1243  
Db 3833 ATGCAAGTAAATATTCGCAAAATGCTTTGATTTGCTATTGTCAGCTCAATGTTGTA 3892  
QY 1244 TyrAsnValIysAlaSerGlySerSerArgArgArgArgSerIleGlnAsnGlnAla 1263  
Db 3893 TATAATGGAAGGCTCTGGGCTCTTGAAGAACCAAGATCTATCCAAATCAAGAGCC 3952  
QY 1264 PheAspLeuAspValAlaValIysGlnAsnIysAspAspLeuAsnIleValAspLeuAsn 1283  
Db 3953 TTTGATTTAGATGTTGCTGTAAGAAATAAAGATGATCTCANTCANTGATTTGAT 4012  
QY 1284 ValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGlnValAsnLeu 1303  
Db 4013 GTGTGTCACAGCTTTCCGGGCCCGGGTGAAGTGGCATGCTCTTATGGAAGTTAACCTA 4072  
QY 1304 LeuSerGlyPheMetValProSerGlnAlaIleSerIleSerGlyIleThrValIlyLeuVal 1323  
Db 4073 TTAAGTGGCTTTATGTCCTTCAGAACCAATTTCTGAGCGAGACAGTGAAGAAAGTG 4132  
QY 1324 GlnTyrAspHisGlyIysLeuAsnLeuTyrLeuAspSerValAsnGlnThrGlnPheCys 1343  
Db 4133 GAATATGATCATGGAACCACTCACTTATTAATTTCTGTAATGAACCACTTTTGT 4192  
QY 1344 ValAsnIleProAlaValArgAsnPheIysValSerAsnThrGlnAspAlaSerValSer 1363  
Db 4193 GTTAATATTCCTGCTGTGAGAACTTTAAAGTTTCAATATCCAAAGATGCTTCAGTGCC 4252  
QY 1364 IleValAspIlyTyrGlnProArgArgGlnAlaValAspSerTyrAsnSerGlnValIys 1383  
Db 4253 ATAGTGATTAATCAAGACCAAGACAGACGGCGGAGAAATTCACACTGGAAGTGAG 4312  
QY 1384 LeuSerSerCysAspLeuCysSerAspValGlnGlyCysArgProCysGluAspGlyAla 1403  
Db 4313 CTGTCCTCTGTCAGCTTTCAGATGATCTCCAGGCTCCCTCTTGAGAGATGAGACT 4372  
QY 1404 SerGlySerHisHisHisSerSerValIlePheIlePheCysPheIysLeuLeuTyrPhe 1423  
Db 4373 TCAGGCTCCCATCATCACTTTCAGTCAATTTTATTTCTGTTCAAGCTTCTGACTTT 4432  
QY 1424 MetGlnLeuTyrPhe 1428  
Db 4433 ATGGAACCTTGGCTG 4447

RESULT 6  
AAD49435  
ID AAD49435 standard; DNA; 5883 BP.  
XX AAD49435;  
AC  
XX  
XX  
XX 24-MAR-2003 (first entry)  
DT  
XX  
XX Human r150 DNA #1.  
DE  
XX  
XX Human; GPI-anchored TGF-beta1 binding protein; r150 protein; cancer;  
KW glycosylphosphatidyl inositol; transforming growth factor-beta1;  
KW therapy; gene; de.  
XX  
OS Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FH 113..4450  
FT CDS  
FT /tag=a  
FT /product="Human r150 protein #2"  
FT /transl\_except="(pos:2219..2221, aa:Xaa)  
FT /note="Xaa corresponds to Ser, Tyr"  
XX  
XX W0200285942-A2.  
XX  
XX 31-OCT-2002.  
XX  
XX 24-APR-2002; 2002WO-CA00560.  
XX

PR 24-APR-2001; 2001US-285713P.  
PR 14-FEB-2002; 2002US-356163P.  
XX  
XX (UYMC-) UNIV MCGILL.  
PA  
XX  
XX Philip A, Tam B;  
XX  
XX WPI; 2003-093100/08.  
DR P-PSDB; AAE32013.  
XX  
XX Novel transforming growth factor (TGF)-beta 1 binding reagent which  
PT comprises r150 protein which acts as accessory receptor of TGF-beta,  
PT useful for negatively modulating TGF-beta activity, and thus for  
PT treating cancer  
XX  
XX Claim 7; Page 100-103; 127pp; English.  
PS  
XX  
XX The invention relates to novel transforming growth factor (TGF)-beta1  
CC binding reagent which comprises a glycosylphosphatidyl inositol (GPI)-  
CC anchored TGF-beta1 binding protein referred to as r150 which acts as  
CC an accessory receptor of TGF-beta. The invention is used for negatively  
CC modulating TGF-beta activity, and thus for treating conditions  
CC characterised by overproduction of TGF-beta, such as cancer. Antisense  
CC molecules of the invention are used for increasing TGF-beta availability  
CC and increase graft success. The present sequence is human r150 protein  
CC encoding DNA.  
XX  
SO Sequence 5883 BP; 1696 A; 1173 C; 1241 G; 1772 T; 1 other;  
Alignment Scores:  
Pred. No.: 0 Length: 5883  
Score: 7318.50 Matches: 1426  
Percent Similarity: 98.75% Conservative: 1  
Best local Similarity: 98.69% Mismatches: 1  
Query Match: 25 Indels: 17  
DB: Gaps: 1

US-10-020-095-4 (1-1428) x AAD49435 (1-5883)  
QY 1 MetGlnGlyProProLeuLeuThrAlaAlaHisLeuLeuCysValCysThrAlaAlaLeu 20  
Db 113 ATGAGGGGCCACCGCTCTGACCGCGCCACCTCTGCTGCTGTCACCGCCGCTG 172  
QY 21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGlyIly 40  
Db 173 GCCGTGGCTCCCGGCTCGCTTCTGTCAGACGCCCAAGGANTCATCAAGCCCGAGGA 232  
QY 41 AsnValThrIleGlyValGlnLeuLeuGlnHisCysProSerGlnValThrValAla 60  
Db 233 AATGTACTATTGGGGTGGAGCTTCTGGAACTGCTCCACAGGTGACTGTGAAGGCG 292  
QY 61 GlnLeuLeuIysThrAlaSerAsnLeuThrValSerValLeuGlnIleGlnIlyValPhe 80  
Db 293 GAGCTGCTCAAGACAGCATCAACCTCAGTCTCTCTCTGGAAGCAGAAAGATCTTT 352  
QY 81 GlnIysGlySerPheIysThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGln 100  
Db 353 GAAGAAAGCTCTTTTAAGACACTTACTCTTCATCACTACTCTGAACAGTGCAGATGAG 412  
QY 101 IleTyrGlnLeuArgValIleThrGlyArgThrGlnAspGlnIleLeuPheSerAsnSerThr 120  
Db 413 ATTATGAGCTACGTGTAAACCGGACGACCAAGATGAGATTTATTTCTTAATATGATACC 472  
QY 121 ArgLeuSerPheGlnThrIlyArgIleSerValPheIleGlnThrAspIlyAlaLeuTyr 140  
Db 473 CGCTTATCATTTAGACCAAGAGAAATATCTGTCTTCAATTAACAGCAAGGCTTATAC 532  
QY 141 LysProIysGlnGlnValIysPheArgIleValThrLeuPheSerAspPheIysProTyr 160  
Db 533 AAGCCAAAGCAAGAGAGATTTCCGATTTGATCACTCTTCCAGATTTTAAGCTTAC 592  
QY 161 LysThrSerLeuAsnIleLeuIleIysAspProIysSerAsnLeuIleGlnIleThrLeu 180

Db 593 AAAACCTCTTAAACATCTCATTAAGACCCCAAAATTAATGATCCAAAGTGGTGG 652  
Qy 181 SerGlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIle 200  
Db 653 TCACAAACAAGATGATTTGGAGTCATTTCCAAACCTTTTCAAGTATCTTCCATCCAAAT 712  
Qy 201 LeuGlyAspTrpSerIleGlnValGlnValAsnAspGlnThrTyrrGlnSerPheGln 220  
Db 713 CTGGGAGATGGTCTATTCAGTTCAAGTGAATGACACAGACATATATCATCATTTTCAG 772  
Qy 221 ValSerGlnTyrValLeuProLysPheGlnValThrLeuGlnThrProLeuTyrCysSer 240  
Db 773 GTTTCAGATATGATATTAACCAAAATTTGAAGTCACTTTCAGACACCATTAATATGTTCT 832  
Qy 241 MetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProVal 260  
Db 833 ATGAAATTCATAGCATTTAAATGATACCATCAACGCAAGATATACATATGGAGGCCAGTG 892  
Qy 261 LysGlyAspValThrLeuThrPheLeuProLeuSerPheTrpGlyLysLysAsnIle 280  
Db 893 AAAGAGACGTAAACGCTTACATTTTACCTTATCTTTGGGGAAGAAGAAATATTT 952  
Qy 281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGlnLysMetLys 300  
Db 953 ACAAAAACATTAAGATTAATGATCTGCAAACCTTCTTTTAATGATGAAGAGATGAA 1012  
Qy 301 AsnValMetAspSerSerAsnGlyLeuSerGlnTyrLeuAspLeuSerProGlyPro 320  
Db 1013 AATGTATATGATTTCTTCAATGACCTTTGTGAATACCTGATCTATCTTCCCTGAGACA 1072  
Qy 321 ValGlnIleLeuThrThrValThrGlnSerValThrGlyIleSerArgAsnValSerThr 340  
Db 1073 GTAGAAATTTTAAACACAGTGAACAGATCAAGTATTTCAAGAAATGTAAGCACT 1132  
Qy 341 AsnValPhePheLysGlnHisAspTyrIleIleGlnPheAspTyrThrThrValLeu 360  
Db 1133 AATGGTCTTCAAGACATATATCAATCATTTGATTTGATATATCTGCTGCTGG 1192  
Qy 361 LysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspLysAsnGlnLeu 380  
Db 1193 AAGCCATCTCTCAACTTCAACGACTGGAAGTAACTCGTGCATGCAACCAACG 1252  
Qy 381 ThrLeuGlnGlnArgArgAsnAsnValValIleThrValThrGlnArgAsnTyrThrGln 400  
Db 1253 ACTCTTGAAGAAGAAGAAATATATGATCATACATGACACAGAGAACTATCTGAG 1312  
Qy 401 TyrTrpSerGlySerAsnSerGlyAsnGlnLysMetGlnAlaValGlnLysIleAsnTyr 420  
Db 1313 TACTGAGCGGATCTPAACAGTGAATCAGAAATGGAAGCTGTTCAAGAAATTAATATAT 1372  
Qy 421 ThrValProGlnSerGlyThrPheLysIleGlnPheProIleLeuGlnLysSerSerGln 440  
Db 1373 ACTGTCCCAAGTGAATCTTTAAGATGAAATTCCTCAATCTGAGAGATTCACAGTGG 1432  
Qy 441 LeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSerLeuPhe 460  
Db 1433 CTACAGTTGAAGCCATATTTCTTGTGATGAAGATGACATGGCAGTTCAATGCTGTTT 1492  
Qy 461 LysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGlnLysIleLysValGly 480  
Db 1493 AAGTCTCCATGATGACATACATCCAACTAAACCAAGAGATGAAATTAAGGGGGA 1552  
Qy 481 SerProPheGlnLeuValValSerGlyAsnLysArgLeuLysGlnLeuSerTyrMetVal 500  
Db 1553 TCGCCTTTGAGTGTGTGTGTGTGCAAAACGATTAAGAGATTAAGCTATATGATGTA 1612  
Qy 501 ValSerArgGlyGlnLeuValAlaValGlyLysGlnAsnSerThrMetPheSerLeuThr 520  
Db 1613 GATTCAGGGGACAGTGTGTGTGTGTGTGCAAAATTCACAAATGTTCTCTTTAAACA 1672  
Qy 521 ProGlnAsnSerTyrThrProLysValAlaCysValIleValTyrTyrIleGlnLysAspGly 540  
Db 1673 CCAGAAATTTCTGACCTCAAAAGCCTGTGTAATGTTGTAATATTAATGAAGATGATGG 1732

Qy 541 GlnIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560  
Db 1733 GAATTTATAGGATGATCTTAAATTTCTGTGACCTTGTGTTTAAAAATATGATTAAG 1792  
Qy 561 LeuTyrTrpSerLysValLysAlaGlnProSerGlnLysValSerLeuArgIleSerVal 580  
Db 1793 CTATATTTGAGATGAAGTGAAGCTGACCATCTGAGAAAGTCTCTTATGATCTCTGG 1852  
Qy 581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600  
Db 1853 ACACAGCCTGATCTCATGTTGGATTTGATGCTGTGCAAAAGTGTATCTGATGAT 1912  
Qy 601 AlaSerAsnAspIleThrMetGlnLysAsnValValHisGlnLeuGlnLeuTyrAsnThrGly 620  
Db 1913 GCCTTAATGATATTTACATGGAATATGTGTGCTCAAGTGTGAACCTTTATTAACACAGA 1972  
Qy 621 TyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGlnCysGlyLeuTrpVal 640  
Db 1973 TATATTTAAGCATGTTCATGATTTCTTTGACGCTTTCAGGAATGTGACCTCGGGTA 2032  
Qy 641 LeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyrAspAsnAlaGlnTyr 660  
Db 2033 TTGACAGATGCAAACTCAGCAAGGATTAATTTGATGCTGTTTATGACAAATGCAAAATAT 2092  
Qy 661 AlaGlnArgPheMetGlnGlnLysAsnGlnGlyHisIleValAspIleHisAspPheSerLeu 680  
Db 2093 GCTGAGAGTTTATGAGAAATATGAAGACATATTTGATGATATTCATGACCTTTCTTGG 2152  
Qy 681 GlySerSerProHisValArgLysHisAspProGlnTyrTrpIleTrpLeuAspThrAsn 700  
Db 2153 GGTACAGATCCACATGTCGAAAGCATTTTCCAGAGACTTGGATTTGGGTAGACACCAAC 2212  
Qy 701 MetGlyTyrArgIleTyrGlnGlnPheGlnValThrValProAspSerIleHisSerTrp 720  
Db 2213 ATGGGTTTCAGGATTTACCAAGAAATTTGAAGTACGTACGATTTATCATCATCTTGG 2272  
Qy 721 ValAlaThrGlyPheValIleSerGlnLysAsnGlyLeuGlyLeuThrThrThrProVal 740  
Db 2273 GTGGCTACGTGTTTGTATCTCTGAGACCTGGGATCTTGGATCAACATCACTCCAGTG 2332  
Qy 741 GlnLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArg 760  
Db 2333 GAGCTCCAAAGCCTTCCAAACATTTTCAATTTTGTGAATCTTCCATCTCTGTATCAGA 2392  
Qy 761 GlyGlnGlnPheAlaLeuGlnIleThrIlePheAsnTyrLeuLysAspAlaThrGlnVal 780  
Db 2393 GGTGAAGATTTGCTTGGAAATATCAATATTTCAATTTTGAAGATGCCACTGAGGT 2452  
Qy 781 LysValIleIleGlnLysSerAspLysPheAspIleLeuMetThrSerSerGlnIleAsn 800  
Db 2453 AAGGTAAATCATGAAAGAAAGTGAACAAATTTGATATTTCAATGACTTCAAAATTAAT 2512  
Qy 801 AlaThrGlyHisGlnGlnIleThrLeuLeuValProSerGlnLysAspGlyAlaThrValLeuPhe 820  
Db 2513 GCCACAGGCCACAGACACATCTTGTGTTCCCATGAGATGGGCAACGTCTTTT 2572  
Qy 821 ProIleArgProThrHisLeuGlyGlnIleProIleThrValThrAlaLeuSerProThr 840  
Db 2573 CCCATCAGGCCAACATCTGGAGAAATTTCTATCAACAGTCAACGCTTTCAACCCACT 2632  
Qy 841 AlaSerAspAlaValThrGlnMetIleLeuValLysAlaGlnGlyIleGlnLysSerTyr 860  
Db 2633 GCTTGTGATGCTGTACCAAGATGATTTTATGAAGGCTGAAGGAATGAAATTAATCATAT 2692  
Qy 861 SerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeu 880  
Db 2693 TCACATCATCTTATTTAGATCTGACATGACATAGGCTACACAGATACCTGAAAACTTTG 2752  
Qy 881 SerPheSerPheProAsnThrValThrGlySerGlnArgValGlnIleThrAlaIle 900  
Db 2753 AGTTTCTCATTTTCTCTTAATATACAGTGAAGTGAAGGATTAATGATCACTCAATTT 2812

QY 901 GlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGly 920  
 Db 2813 GGAGATGTTCTTGGCTTCCTTCATCAATGGCTTAGCGCTCATGATTCGGATCCTTAATGGC 2872  
 QY 921 CysGlyGluGlnAsnMetIleAsnPheHlaProAsnIleTyrIleLeuAspTyrLeuThr 940  
 Db 2873 TGTGGTGAACAGAACATGATTAATTTGCTCCAAATATTACATTTTGGATTATTCAGCT 2932  
 QY 941 LysIleLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPheMetArgGlnGly 960  
 Db 2933 AAAAAGAAACAACACACAGATTAATTTGAAGAAAAGCTCTTTCATTTATAGCCAAAGT 2992  
 QY 961 TyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyr 980  
 Db 2993 TACAGAGAGAACTCTCTATACAGAGGAAATGGCTCTTCAAGCTTTGGGAATAT 3052  
 QY 981 AspProSerGlySerThrTyrLeuSerAlaPheValLeuArgCysPheLeuGluAlaAsp 1000  
 Db 3053 GACCTTCTGGAGACATTGGTGTGACGCTTTGTTTAAAGATGTTCTTGAAGCCGAT 3112  
 QY 1001 ProTyrIleAspIleAspGlnAsnValLeuHlaArgThrTyrThrTyrLeuLysGlyHis 1020  
 Db 3113 CTTTACATAGATTTGATCAGAAATGTGTACACAGAAATACACTGGCTTAAAGACAT 3172  
 QY 1021 GlnLysSerAsnGlyGluPheTyrAspProGlyArgValIleHisSerGluLeuGlnGly 1040  
 Db 3173 CAGAAATCCAACGGTGAATTTTGGATCCAGAAAGATGATTCAATGAGACTTCAAGT 3232  
 QY 1041 GlnLysLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArg 1060  
 Db 3233 GGCATATMAAGTCCAGTAACTTAACAGCTTAATGTGAACCTTCTCTCGGGATATAGA 3292  
 QY 1061 LysTyrGlnProAsnIleAspValGlnLysSerIleHisPheLeuGluSerGluPheSer 1080  
 Db 3293 AAGATTCAGCTTAACATTTGATGTGCAGAGCTTATCCATTTTGGAGCTGAAATTCAGT 3352  
 QY 1081 ArgGlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGly 1100  
 Db 3353 AGAGGAATTTTCAACAATTAATTAATCTAGCCCTTAATTAATTAATTAATTAATTAAT 3412  
 QY 1101 SerProLysAlaLysGluAlaLeuAsnMetLeuThrTyrPheAlaGluGlnGlnGly 1120  
 Db 3413 AGTCCTAAAGGAGGAGGAGCTTTGAATATGTGACTTGAAGAGAGAGGAGGAGGAGG 3472  
 QY 1121 MetGlnPheTyrValSerSerGluSerLysLeuSerAspSerTyrGlnProArgSerLeu 1140  
 Db 3473 ATGCATTTCTGGGTGTCATCAGAGTCCAACTTCTGACCTCTGGACCCACGCTCCCTG 3532  
 QY 1141 AspIleGluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGlu 1160  
 Db 3533 GATATTGAAGTTGACAGCTTATGACCTGCTCACACTTCAACATTTCAAGCTTCTGAG 3592  
 QY 1161 GlyIleProIleMetLysTyrPheLysSerArgGlnArgAsnSerLeuGlyGlyPheAlaSer 1180  
 Db 3593 GGAATCCCAATTATGAGGTGCTAAGCAGGCAAAATTAACCTTGGGGTGTTCATCT 3652  
 QY 1181 ThrGlnAspTyrThrValAlaLeuLysAlaLeuSerGluPheAlaAlaLeuMetAsnThr 1200  
 Db 3653 ACTGATGATACCTGCTGCTGCTTAAAGCTGCTGCTGATTTGAGCCCTTAATAGAACCA 3712  
 QY 1201 GluArgThrAsnIleGlnValThrValThrGlyProSerSerProSerPro----- 1217  
 Db 3713 GAAAGGACAAATATCCAAAGTACCGGTGACGGGGCTTACCAAGTCCGTAAAGTTT 3772  
 QY 1218 -----LeuAlaValValGlnPro 1223  
 Db 3773 CTGATTGACACACAACCGCTTACTCTTCAGACAGACAGACTGCTGTGTACAGCCA 3832  
 QY 1224 MetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValVal 1243  
 Db 3833 ATGGCAGTTAATATTTCCGAAATGGTTTGGATTGTCTATTTGTTCAGCTCAAGTGTGTA 3892  
 QY 1244 TyrAsnValLysAlaSerGlySerSerArgArgArgSerIleGlnAsnGlnGluAla 1263

Db 3893 TATATGTGAAGCTTCTGGGCTCTTACAGAGAGGAAATCTATCCAAATCAAGAGCC 3952  
 QY 1264 PheAspLeuAspValAlaValLysGluAsnLysAspAspLeuAsnHisValAspLeuAsn 1283  
 Db 3953 TTGATTTAATGTTGCTGTAAAGAAATTAAGATGATCTCAATCATGTGGATTTCAT 4012  
 QY 1284 ValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluValAsnLeu 1303  
 Db 4013 GTGTGATCAAGCTTTTGGGCGCGGTAGAGATGGCATGGCTCTTATGAAAGTTAACTTA 4072  
 QY 1304 LeuSerGlyPheMetValProSerGluAlaIleSerLeuSerGluThrValLysLysVal 1323  
 Db 4073 TTAAGTGCTTTATGTGGCTTCAAGAACAAATTTCTGTGACGACAGACAGTGAAGAAAGT 4132  
 QY 1324 GluTyrAspHisGlyLysLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCys 1343  
 Db 4133 GAATATATCATGGAACACCAACCTCTATTTGATTTCTTAAATGAACCCAGTTTGT 4192  
 QY 1344 ValAsnIleProAlaValAlaArgAsnPheLysValSerAsnThrGlnAspAlaSerValSer 1363  
 Db 4193 GTTATATTCCTGCTGTGAGAACTTTAAAGTTTCAATCCCAAGATGCTTCAGTGTCC 4252  
 QY 1364 IleValAspTyrTyrGluProArgArgGlnAlaValArgSerTyrAsnSerGluValLys 1383  
 Db 4253 ATAGTGATTAATCATGAGCCAGAGAGCAGCGGTGAGAGTTACAACTCTGAAGTCAG 4312  
 QY 1384 LeuSerSerCysAspLeuCysSerAspValGlnGlyCysArgProCysGluAspGlyAla 1403  
 Db 4313 CTGCTCTCTGTGACCTTTGACATGATGTCCAGGGCTGCGCTCTGTGAGATGAGACT 4372  
 QY 1404 SerGlySerHisHisSerSerValIlePheIlePheCysPheLysLeuTyrPhe 1423  
 Db 4373 TCAGGCTCCCATCATCATCTTCAATGATTTTATTTTCTGTTTCAAGCTTCTGATTTT 4432  
 QY 1424 MetGluLeuTyrLeu 1428  
 Db 4433 ATGCACTTGGCTG 4447  
 RESULT 7  
 AAL49816 standard; cDNA; 4335 BP.  
 ID AAL49816;  
 AC AAL49816;  
 XX  
 DT 04-DEC-2002 (first entry)  
 XX  
 DE Human platelet alloantigen Govb coding sequence.  
 XX  
 KW Human; platelet alloantigen; Govb; single nucleotide polymorphism;  
 KW SNP; diallelic; bleeding disorder; post-transfusion purpura;  
 KW post-transfusion platelet refractoriness; haemostatic; vaccine;  
 KW neonatal alloimmune thrombocytopenia; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..4335  
 FT /tag= a  
 FT /product= "Govb"  
 FT /partial  
 FT /note= "no stop codon"  
 FT replace(2108,A)  
 FT /tag= b  
 XX  
 PN MO200270738-A2.  
 XX  
 PD 12-SEP-2002.  
 XX  
 PF 07-MAR-2002; 2002NC-CAD00291.  
 XX  
 PR 07-MAR-2001; 2001US-273941P.  
 XX

PA (SCHU/) SCHU A.  
 XX Schu A, Ouwehand W;  
 XX MPI: 2002-713460/77.  
 DR P-PSDB; AAO19373.  
 XX  
 PT New isolated oligonucleotide binding to a region of CD109 nucleic acid  
 PT having a single nucleotide polymorphism that distinguishes a Gova  
 PT and/or Gova allele, useful for treating blood disorders e.g. autoimmune  
 PT thrombocytopenia  
 XX  
 PS Claim 7, Page 35-41; 69pp; English.  
 XX  
 CC The present invention relates to a sequence capable of binding  
 CC specifically to a CD109 nucleic acid which has a single nucleotide  
 CC polymorphism that distinguishes the Gova and Gova alleles. Detection of  
 CC the Gova genotype is useful for detecting whether the subject has or is at  
 CC risk of a blood disease, disorder or abnormal physical state, such as  
 CC bleeding, or increased risk of bleeding, due to autoimmune destruction of  
 CC blood platelets, e.g., post-transfusion purpura, post-transfusion  
 CC platelet refractoriness or neonatal autoimmune thrombocytopenia. The  
 CC nucleic acid and polypeptide are useful for Gova genotyping or phenotyping  
 CC individuals. The present sequence is the Gova coding sequence.  
 XX  
 SQ Sequence 4335 BP; 1295 A; 838 C; 922 G; 1280 T; 0 other;

## Alignment Scores:

Pred. No.: 0 Length: 4335  
 Score: 7317.50 Matches: 1426  
 Percent Similarity: 98.75% Conservative: 1  
 Best Local Similarity: 98.69% Mismatches: 1  
 Query Match: 99.58% Indels: 17  
 DB: 24 Gaps: 1

US-10-020-095-4 (1-1428) x AAL49816 (1-4335)

QY 1 MetGInGlyProProLeuLeuThraAlaAlaHsLeuLeuCyValCyThrAlaAlaLeu 20  
 DB 1 ATGCGAGGCGCCACCGCTCTGACCGCGGCCACCTCTGCTGCTGCGACCGCGCGCTG 60  
 QY 21 AlaValAlaProGlyProAlaGlyPheLeuValThraAlaProGlyIleIleArgProGly 40  
 DB 61 GCGGTGGCTCCGCGGCTCGGTTCTGTGACAGCCCGAGGATCATCAGCCCGAGGA 120  
 QY 41 AenValThraIleGlyValGluLeuLeuGluHsCyProSerGlnValThraValAla 60  
 DB 121 AATGTGACTAATGGGGTGGAGCTTCTGGAACACTGCCCTTCAAGGTGACTGTGAAGCG 180  
 QY 61 GluLeuLeuThraAlaSerAsnLeuThraValSerValLeuGluAlaGluGlyValPhe 80  
 DB 181 GAGGTGCTCAAGACACATCACTCACTGCTCTGCTGAGACAAAGAGAGCTTT 240  
 QY 81 GluLysGlySerPheLeuThraLeuThraProSerLeuProLeuAsnSerAlaAspGlu 100  
 DB 241 GAAAAAGGCTCTTTAAGACACTTACTTCCATCACTACCTCTGAACAGTCAAGTAG 300  
 QY 101 IleTyrgluLeuArgValThraGlyArgThraAspGluIleLeuPheSerAsnSerThr 120  
 DB 301 ATTTATGACCTACGCTTAACCGAGCTACCCAGATGAGATTTATTTCTTAATAGTACC 360  
 QY 121 ArgLeuSerPheGluThraIleValIleSerValPheIleGlnThraAspLysValLeuTy 140  
 DB 361 CGCTTAATCAATTGAGACCAAGAAATATATGCTTCAATCAACAGACAAAGGCTTATAC 420  
 QY 141 LysProLysGlnGluValLysPheArgIleValThraLeuPheSerAspPheLysProTy 160  
 DB 421 AAGCCAAACCAAGAGTGAAGTTTGCATTGTTCACCTCTCTCAGATTTTAAGCTTAC 480  
 QY 161 LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGlnTrpLeu 180  
 DB 481 AAAACCTTTTAAACATCTCATTAAGAGACCCCAATCAAAATTTGATTCACAGTGGTGG 540

QY 181 SerGlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIle 200  
 DB 541 TCACAAACAAGTGAATCTTGAGAGTCAATTCACAAACCTTTTGCATATCTTCCATCATA 600  
 QY 201 LeuGlyAspTrpSerIleGlnValGlnValAsnAspGlnThrTyTyrglnSerPheGln 220  
 DB 601 CTGGTGACTGCTCATTAATCAAGTTCACAGTGAATGACCAAGACATATTAATCAATTCAG 660  
 QY 221 ValSerGluTyValLeuProLysPheGluValThraLeuGlnThrProLeuTyCySer 240  
 DB 661 GTTTAGAAATATGATTTCCAAATTTCAAGTGAATGACCTTTGAGACACATTAATTTGTTCT 720  
 QY 241 MetAsnSerLysHisLeuAsnGlyThrIleThraAlaTyTyrglyLysProVal 260  
 DB 721 ATGAATTTCAAGCATTTAATATGATGATCAACGGAAGTATGATATGGAAGCAGTG 780  
 QY 261 LysGlyAspValThraLeuThraPheLeuProLeuSerPheTrpGlyLysLysAsnIle 280  
 DB 781 AAGAGAGACGTACAGCTTACATTTTACCTTTATCTTTGGGGAAGAGAAATATAT 840  
 QY 281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluGluMetLys 300  
 DB 841 ACAAACAATTTAAGATTAATGATGATGCAACATCTCTTTAATGATGAAGAGTGA 900  
 QY 301 AenValMetAspSerSerAsnGlyLeuSerGluTyriLeuAspLeuSerProGlyPro 320  
 DB 901 AATGTAAATGATTTCTTCAATGAGACTTCTGAAATCCGTGATCATCTTCCCTGAGACA 960  
 QY 321 ValGluIleLeuThraThraValThraGluSerValThraGlyIleSerArgAsnValSerThr 340  
 DB 961 GTAGAAATTTTAACACAGTACAGAAATCAATTCAGGTATTTCAAGAAATGTAAGCACT 1020  
 QY 341 AenValPhePheLysGlnHisAspTyriIleGluPhePheAspTyriThraValLeu 360  
 DB 1021 AATGTGCTTCAAGCAACATGATTAATCAATGATGATTTTGTGATTAATCACTGCTCTTG 1080  
 QY 361 LysProSerLeuAsnPheThraAlaThraValLysValThraArgAlaAspGlyAsnGlnLeu 380  
 DB 1081 AAGCATCTCTCACTTCACTTCAACGCACTGTGAAGTAACTCGTGTGATGAGCAACCACTG 1140  
 QY 381 ThrLeuGluGluArgArgAsnAsnValIleThraValThraGlnArgAsnTyriThraGlu 400  
 DB 1141 ACTCTTGAAGAAAGAAATTAATGATGATCAATCAAGAGACACAGAAACTATCTAG 1200  
 QY 401 TyrTrpSerGlySerAsnSerGlyAsnGlnLysMetGluAlaValGlnLysIleAsnTy 420  
 DB 1201 TACTGAGCGGATCTCAACAGTGAATCAAGAAATGGAAGCTGTTCAAGAAATTAATAT 1260  
 QY 421 ThrValProGlnSerGlyThraPheLysIleGluPheProIleLeuGluAspSerSerGlu 440  
 DB 1261 ACTGTCCCCCAAGTGAACCTTTTAAGTTAATTTCCAAATCTCGAGAGATTTCCAGTAG 1320  
 QY 441 LeuGlnLeuLysAlaTyriPheLeuGlySerLysSerSerMetAlaValHisSerLeuPhe 460  
 DB 1321 CTACAGTGAAGGCTTATTTCTTCTGTGTAAGTAGACATGGCAGTTCAATAGCTGT 1380  
 QY 461 LysSerProSerLysThrTyriIleGlnLeuLysThraArgAspGluAsnIleLysValGly 480  
 DB 1381 AAGTCTCTTAATGAACATCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 1440  
 QY 481 SerProPheGluLeuValIleSerGlyAsnLysArgLeuLysGluLeuSerTyMetVal 500  
 DB 1441 TCGCTTTTGAAGTGGTGTGATGAGCAACAAAGATTTGAAGAGCTTAAGCTATATG 1500  
 QY 501 ValSerArgGlyGlnLeuValAlaValGlyLysGlnAsnSerThraMetPheSerLeuThr 520  
 DB 1501 GTATCAAGGAGACAGTGTGTGAGAGAAACAAATTCACAAATGCTTCTTTAACA 1560  
 QY 521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyTyriIleGluAspAspGly 540  
 DB 1561 CCAGAAATTTCTTGACCTCCAAAGCCGTGTAAATGTGTATTTAATGAAGATGATGG 1620  
 QY 541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560

Db 1621 GAAATTAATGAAGTGGTCTTAAAAATTCCTGCTGCTGTTTAAAAAATGAAGTAAAG 1680  
Qy LeuTYRTPSerLysValIleValAGIupProSerGIuLysValSerLeuArgIleSerVal 580  
Db 1681 CTATATTTGGAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 1740  
Qy ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600  
Db 1741 AACAGAGCTGACCTCATAGTTGGGATGTAGCTGTTGACAAAGTAAGTGAATCTGATGAT 1800  
Qy AlaSerAsnAspIleThrMetGluAsnValValHisGluLeuGluLeuTYRAsnThrGly 620  
Db 1801 GCCCTTAATGATTAATTAACAATGAAATGTGTCCATGATGTGAACCTTAATAACAAGCA 1860  
Qy TYRTrpLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTrpVal 640  
Db 1861 TATTTATTAAGCATTTATGATGATTTCTTTGCAAGCTTTCCAGAAATGTGACTGTGGGTA 1920  
Qy LeuThrAspAlaAsnLeuThrLysAspTYRILEAspGlyValTYRAspAsnAlaGluTYR 660  
Db 1921 TTGACAGATGCAAACTCAGCAAGATTAATATGATGTTGTTATGACAAATGCAAGATAT 1980  
Qy AlaGluAspPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680  
Db 1981 GCTGAGAGCTTTATGAGAAATGAAGACATATTTGATGATATTCATGACCTTTTCTTTG 2040  
Qy GlySerSerProHisValArgLysHisAspPheGluTrpIleTrpLeuAspThrAsn 700  
Db 2041 GGTACACATCCACATGCTCGAAAGCATTTCCAGAGACTTGATTTGGCTAGACCCAC 2100  
Qy MetGlyTYRArgIleTYRglnGluPheGluValThrValProAspSerIleThrSerTrp 720  
Db 2101 ATGGGATCCAGATTTACCAAGAAATTTGAGTAACCTGATTCATCATCTTCTTTGG 2160  
Qy ValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrProVal 740  
Db 2161 GTGGTACTGTTTGTGATCTCTGAGAGCCTGGGCTTGGACCTAACACTACCTCCAGTG 2220  
Qy GluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTYRSerValIleArg 760  
Db 2221 GAGCTCCAAAGCTTCCAAACATTTTTCATTTTGTGAATCTTCCCTACTCTGTATACGA 2280  
Qy GlyGluGluPheAlaLeuGluIleThrIlePheAsnTYRLeuLysAspAlaThrGluVal 780  
Db 2281 GGTGAAGATTTGCTTGGAAATACATATTCATATTTGAAAGATGCCACGAGGTT 2340  
Qy LysValIleIleGlyLysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsn 800  
Db 2341 AAGTAATCATTTGAGAAAGTACAAATTTGATATTTCAATGACTTCAATGAATGAATAAAT 2400  
Qy AlaThrGlyHisGlnGlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPhe 820  
Db 2401 GCCACAGGCCACAGACACCTTCTGTGCTCCAGTGGAGATGGGGCAACTGTTCTTTT 2460  
Qy ProIleArgProThrHisLeuGlyGluIleProIleThrValThrAlaLeuSerProThr 840  
Db 2461 CCACATCAGGCCAACATCTGGAGAAATTCCTATCAACATCAGCTCTTCCACCACT 2520  
Qy AlaSerAspAlaValThrGluMetIleLeuValLysAlaGluGlyIleGlyLysSerTYR 860  
Db 2521 GCTTCTGATGCTGTCACCAACATGATTTTAAAGGCTGAAGAAATGAATAATCATAT 2580  
Qy SerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeu 880  
Db 2581 TCACATATCATTTATTAAGCTGTGACTGACATAGGCTACACAGATACCTTGAATTAATTTG 2640  
Qy SerPheSerPheProAsnThrValThrGlySerGluArgValGlnIleThrAlaIle 900  
Db 2641 AGTTTCATATTTCTCCCTAATACAGTACCTGGCAGTGAAGATTCAGATACCTGCATTT 2700  
Qy GlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTYRgln 920

Db 2701 GGAAGTCTTGGTCTCTCCATCATGATGCTTAAGCTCATTAATGATGATGCTTATGGC 2760  
Qy CysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTYRILEAspTYRLeuThr 940  
Db 2761 TGTGTGAACAGAAACGATTAATTTTGTCTCCAAATATTTACATTTTGGATTAATGACT 2820  
Qy LysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPheMetArgGlnGly 960  
Db 2821 AAAAAGAAACACATGACAGATTAATTTGAAAGAAAAGCTTTTCATTTATGAGGCAAGT 2880  
Qy TYRglnArgGluLeuLeuTYRglnArgGluAspGlySerPheSerAlaPheGlyAsnTYR 980  
Db 2881 TACCAGAGAAACTTCTCATACAGAGGAAAGAGGCTTTTCAGTGTGCTTTGGAAATTAAT 2940  
Qy AspProSerGlySerThrTrpLeuSerAlaPheValLeuArgCysPheLeuGluAlaAsp 1000  
Db 2941 GACCCCTTCTGGAGCACTTGTTGTGACGCTTTGTTTAAAGATGTTTCCCTGAAGCCGAT 3000  
Qy ProTYRILEAspIleAspGlnAsnValLeuHisArgTYRTrpTrpLeuLysGlyHis 1020  
Db 3001 CCTTACATAGATATTGATCAGAAATGTGTACACAGAAACATACACTTGGCTTAAGGACAT 3060  
Qy GlnLysSerAsnGlyGluPheTrpAspProGlyArgValIleHisSerGluLeuGlnGly 1040  
Db 3061 CAGAAATCCAAAGGTGATTTTGGGATCCAGAGAGGTGATTCATAGTGTAGCTTCAAGT 3120  
Qy GlyAsnLysSerProValThrLeuThrAlaTYRILEValThrSerLeuLeuGlyTYRArg 1060  
Db 3121 GGCATTAAGTAAGTCCAGTAACACTTACAGCTTAATTTGTATCTTCTCCGGGATATAGA 3180  
Qy LysTYRglnProAsnIleAspValGlnGluSerIleHisPheLeuGluSerGluPheSer 1080  
Db 3181 AAGTATCAGGCTTAACATGATGTGCAAGAGTCTATCATATTTTGGAGTGTGAATTCAGT 3240  
Qy ArgGlyIleSerAspAsnTYRTrpLeuAlaLeuIleThrTYRTrpAlaLeuSerSerValGly 1100  
Db 3241 AGAGAAATTTCAACATTAATTAATCTTAAGCTTATTAATCTTAATTAATCTTAATTAAT 3300  
Qy SerProLysAlaLysGluAlaLeuAsnMetLeuThrTrpArgAlaGluGlnGluGly 1120  
Db 3301 AGTCTTAACAGAAAGAGAGCTTTGAATATCTGACTTGAAGACAGAAAGAGTGGC 3360  
Qy MetGlnPheTrpValSerSerGluSerLysLeuSerAspSerTrpGlnProArgSerLeu 1140  
Db 3361 ATCAATATTCGGGTGCATCAGAGTCCAAACTTTCGACTTCCAGGACGACCTCCCTG 3420  
Qy AspIleGluValAlaAlaTYRAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGlu 1160  
Db 3421 GATATGAAAGTTCAGCTTAATGACATGCTCTCACACTTCTTAATTAATTTGACTTGTAG 3480  
Qy GlyIleProIleMetArgTrpLeuSerArgGlnArgAsnSerLeuGlyPheAlaSer 1180  
Db 3481 GGAATCCCAATTTATGAGTGTGCTTAAGCAGGCAAGAAATAGCTTGGGTGTTTGCATCT 3540  
Qy ThrGlnAspThrThrValAlaLeuLysValAlaLeuSerGluPheAlaIleLeuMetAsnThr 1200  
Db 3541 ACTCAGAGTACCACTGGCTTTAAAGGCTCTGTGCAATTTGACAGCCCTTAATGAATACA 3600  
Qy GlnArgThrAsnIleGlnValThrValThrGlyProSerSerProSerPro----- 1217  
Db 3601 GAAAGGACAAATATCCAAAGTACCGTGAACGGGGCTTAAGTCAACAACTCTTGAAGTTT 3660  
Qy -----LeuAlaValAlaGlnPro 1223  
Db 3661 CTGATTAACACACAAACCGCTTAATCTTCCAGACAGCAGAGCTTGTGCTGTRACAGCCA 3720  
Qy MetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValVal 1243  
Db 3721 ATGGCAGTTAATATTTCCGGAATATGTTTGAATTTGATTTGATTTGATTTGATTTGATTT 3780  
Qy TYRAsnValIleLysAlaSerGlySerSerArgArgArgSerIleGlnAsnGlnGluAla 1263  
Db 3781 TATTAATGTGAAGGCTTCTGGGCTTCTTAAGAGAGCAAGATCTATCCAAATCAAGAGGCC 3840





QY 161 LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGlnThrLeu 180  
Db 593 AAAACCTCTTTAAACATTCTCATTTAGAGACCCCAATCAAAATTGATCAACAGGTGG 652  
QY 181 SerGlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIle 200  
Db 653 TCACAAACAAAGTGAAGTCTGGAGTCAATTCCTTCCAAACCTTTCAGCTATCTTCCATCCATA 712  
QY 201 LeuGlyAspTrpSerIleGlnValGlnValAsnAspGlnThrLysThrLysGlnSerPheGln 220  
Db 713 CTGGTGCCTGCTCTATTCAAGTTCAAGTGAATGACGACACATATTATCAATTCATTTCAG 772  
QY 221 ValSerGlnLysValLeuProLysPheGlnValThrLeuGlnThrProLeuLysCysSer 240  
Db 773 GTTTCAGAAATATGATTTACCAAAATTTGAAGTCACTTTCAGACACCATTTATTTGTTCT 832  
QY 241 MetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysLysThrLysGlyLysProVal 260  
Db 833 ATGAATTCCTAAGCATTTAAATGCTACATCACGCGCAAGTATACATATGGAGACCAAGTG 892  
QY 261 LysGlyAspValThrLeuThrPheLeuProLysSerPheTrpGlyLysLysValAsnIle 280  
Db 893 AAAGAGACGTACGCTTACATTTTACCTTTATCTTTGCGGAAGAGAAAATATTT 952  
QY 281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGlnGluMetLys 300  
Db 953 ACAAACAACTTTAAGTAAATGAGATCTGCAAACTTCTTTTATGATGAGAAAGATGAAA 1012  
QY 301 AsnValMetAspSerSerAsnGlyLeuSerGlnLysLeuAspLeuSerSerProGlyPro 320  
Db 1013 AATGAATGAGATTTCTTCAAAATGAGACTTCTGAATCTCGATCTATCTTCCCTGGACCA 1072  
QY 321 ValGluIleLeuThrThrValThrGlnSerValThrGlyIleSerArgAsnValSerThr 340  
Db 1073 GTAGAAATTTTAAACACAGTACAGAAATGAGTATTTCAAGAAATGTAAGCACT 1132  
QY 341 AsnValPhePheLysGlnHisAspLysIleIleGlnPheAspLysThrThrValLeu 360  
Db 1133 AATGTTCTTCAAGCAACATGATTACATCATTTGATTTTGAATTAACACTGCTCTG 1192  
QY 361 LysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsnGlnLeu 380  
Db 1193 AAGCATCTCTCAACTTCACAGCCACTGTGAAGGAACTCGTGTGATGCGCAACCACTG 1252  
QY 381 ThrLeuGlnLysArgArgAsnAsnValIleThrValThrGlnArgAsnThrThrGlu 400  
Db 1253 ACTCTGAAGAAAGAAATTAATGATGATCAATACAGTACACAGAGAACTATATCTGAG 1312  
QY 401 TyrTrpSerGlySerAsnSerGlyAsnGlnLysMetGlnAlaValGlnLysIleAsnTyr 420  
Db 1313 TACTGAGCGGATCTTAACAGTGGAAATCAGAAATGGAAGCTGTTCAAGAAATTAATAT 1372  
QY 421 ThrValProGlnSerGlyThrPheLysIleGlnPheProIleLeuGlnAspSerSerGlu 440  
Db 1373 ACTGCCCCCAAGGAACTTTTAAAGATGAATTCCTCAATCTCGAGAGATTTCCAGTGA 1432  
QY 441 LeuGlnLeuLysAlaTyrPheLeuGlnLysSerLysSerSerMetAlaValHisSerLeuPhe 460  
Db 1433 CTACAGTTGAAGGCTTATTTCTTGGTGAAGTAAAGTACATGCGAGTTCATATAGTCTGTT 1492  
QY 461 LysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGlnAsnIleLysValGly 480  
Db 1493 AAGTCTCTAGTAAACATACATCACTCAATAAAACAGAGATGAAGAAATTAAGTGGGA 1552  
QY 481 SerProPheGlnLeuValValSerGlyAsnLysArgLeuLysGlnLeuSerTyrMetVal 500  
Db 1553 TCGCCTTTTGAAGTGGTGTAGTGAACAACAGATGAAGAGTGAAGCTATATAGTGA 1612  
QY 501 ValSerArgGlyGlnLeuValAlaValGlyLysGlnAsnSerThrMetPheSerLeuThr 520  
Db 1613 GTATCCAGGAGACAGTTGGTGGCTGTAGAAAAACAAATTCACAAATGTTCTCTTAAACA 1672

QY 521 ProGlnAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGlnAspAspGly 540  
Db 1673 CCAAGAAATTTCTGACTCCAAAAGCCTGTGATTTGTATTAATTAATGAAGTGAATGG 1732  
QY 541 GlnIleIleSerAspValIleLysPheIleProValGlnLeuValPheLysAsnLysIleLys 560  
Db 1733 GAAATTAAGTATGATGTTCAAAAATTCCTGTCAGCTGTGTTTAAATAAAGATTAAG 1792  
QY 561 LeuTyrTrpSerLysValLysAlaGlnProSerGlnLysValSerLeuArgIleSerVal 580  
Db 1793 CTATATTGAGTAAAGTGAAGAGCTGAACCATCTGAGAAATCTCTCTAGATCTCTGTG 1852  
QY 581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600  
Db 1853 ACACAGCCTGATCTCAAGTGAATGTGATGTGCTGTGACAAAGTGAATTCGATGAAT 1912  
QY 601 AlaSerAsnAspIleThrMetGlnAsnValValHisGlnLeuGlnLeuLysTrpAsnThrGly 620  
Db 1913 GCCTTATATATATATTAATGAAATGTGTCTCAGAGTTGGAACTTTAAACACAGGA 1972  
QY 621 TyrTrpLeuGlyMetPheMetAsnSerPheAlaValPheGlnGlnLysGlyLeuTrpVal 640  
Db 1973 TATTTATTAGGCATGTTTCAATGAATCTTTTGCAGTCTTTCAGGAATCTGGACTGGGTA 2032  
QY 641 LeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyrAspAsnAlaGlnTyr 660  
Db 2033 TTGACAGATCAAACTCACGAAGAGATTATATGATGTGTTATATGACATGACAGAAATAT 2092  
QY 661 AlaGlnArgPheMetGlnLysGlnGlnGlyHisIleValAspIleHisAspPheSerLeu 680  
Db 2093 GCTGAGAGGTTTATGAGGAAATGAAAGACATATGTATATTCATATGACTTTTCTTTG 2152  
QY 681 GlySerSerProHisValArgLysHisPheProGlnThrTrpIleTrpLeuAspTrpAsn 700  
Db 2153 GGTAGCATCCACATGTCCGAAGACATTTTCCAGAGATCTTGATTTGCTAGACACCAAC 2212  
QY 701 MetGlyTyrArgIleTyrGlnGlnPheGlnValThrValProAspSerIleThrSerTyr 720  
Db 2213 ATGGGTTCCAGATTTTAAACAAATTAATGAAGTAACTGATCTGATTCATACCTTCTGG 2272  
QY 721 ValAlaThrGlyPheValIleSerGlnAspLeuGlyLeuGlnLysLeuThrThrProVal 740  
Db 2273 GTGGCTACTGTTTGTGATCTCTGAGACCTGGCTGTGGACTTAACATACTCCAGTG 2332  
QY 741 GlnLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArg 760  
Db 2333 GAGCTCCAGCCTTCCACACATTTTCTTCTTTTGAATTTTCCCTACTCTGTATCAGA 2392  
QY 761 GlyGlnGlnPheAlaLeuGlnIleThrIlePheAsnTyrLeuLysAspAlaThrGlnVal 780  
Db 2393 GGTGAAGAAATTTCTTGGAAATTAACATATTAATTAATTTGAAGTGCACATGAGTT 2452  
QY 781 LysValIleIleGlnLysSerAspLysPheAspIleLeuMetThrSerSerGlnIleAsn 800  
Db 2453 AAGTATCACTTGAAGAAAGTGAACAAATTTGATTTCTATATGACTTCAATGAAGAAAT 2512  
QY 801 AlaThrGlyHisGlnGlnThrLeuLeuValProSerGlnLysAspGlyAlaThrValLeuPhe 820  
Db 2513 GCCACAGGCCACAGCAGACCTTCTGTTCCAGTGAAGATGGGCACTGTTCTTTT 2572  
QY 821 ProIleArgProThrHisLeuGlnGlyGlnIleProIleThrValThrAlaLeuSerProThr 840  
Db 2573 CCATCAAGGCCAACACATCTGGGAAATTCCTATCAAGTCAAGCTCTTCAACCACT 2632  
QY 841 AlaSerAspAlaValThrGlnMetIleLeuValLysAlaGlnGlyIleGlnLysSerTyr 860  
Db 2633 GCTTCTGATCTCTGACCCAGATGATTTTATGTAAGCTGAAGAAATGAAGAAATCATAT 2692  
QY 861 SerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeu 880  
Db 2693 TCACAAATCAATCTTATTAAGCTTGAAGTGAAGGCTTACAGATCCCTGAAAACCTTTG 2752  
QY 881 SerPheSerPheProAsnThrValThrGlySerGlnArgValGlnIleThrAlaIle 900

```

Db      2753  |||||ACCTTCTCATTCTCCCTAATAACAGTACGACGAGAAAGATTCACTGCAAT 2812
Qy      901  |||||GlyAspValLeuGlyProSerIleasnGlyLeuAlaSerLeuIleArgMetProTyrGly 920
Db      2813  |||||GGAGATGTTCTTGTCCTTCCTCAATCAATGCTTACATGATTCAGGAGGCTTATGAC 2872
Qy      921  |||||CysGlyGluGlnAsnMetIleasnPheAlaProAsnIleTyrIleuAspTyrLeuThr 940
Db      2873  |||||TGTGTCGAACAGAACATGATTAATTTGCTCCAAATATTACATTTTGATTTATCTGACT 2932
Qy      941  |||||LysIleValGlnLeuThrAspAsnLeuLysGlyValAlaLeuSerPheMetArgGlnGly 960
Db      2933  |||||AAAAAGAAACAACCTGCACATATATTGAAAGAAAAAGCTCTTCAATTATGAGGCAAGCT 2992
Qy      961  |||||TyrGlnAspGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyr 980
Db      2993  |||||TACAGAGAGAACTTCTTATACAGAGGAGAGAGGCTCTTTCAGTCTTTTGAGAAATTAT 3052
Qy      981  |||||AspProSerGlySerThrTyrLeuSerAlaPheValLeuArgCysPheLeuGluAlaAsp 1000
Db      3053  |||||GACCTTCTGGGAGACCTTGTTGTCAGCTTTGTTTAAGATGTTTCTTGAGCCGAT 3112
Qy      1001  |||||ProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTyrLeuLysGlyHis 1020
Db      3113  |||||CTTTCATATGATATTGATCAGATGTTGTTACACAGAACATACACTTGCTTAAAGACAT 3172
Qy      1021  |||||GlnLysSerAsnGlyGlnPheTyrAspProGlyArgValIleHisSerGluLeuGlnGly 1040
Db      3173  |||||CAGAAATCCACGCGGATTTGGATCCAGAGAGGATTCATATGATGAGCTTCAAGCT 3232
Qy      1041  |||||GlyAsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuGlnGlyTyrArg 1060
Db      3233  |||||GGCAATTAAAGTCCAGTAACTTACAGGCTATATTGTAATCTTCTCTGGGATTTAGA 3292
Qy      1061  |||||LysTyrGlnProAsnIleAspValGlnGlnSerIleHisPheLeuGlnSerGluPheSer 1080
Db      3293  |||||AAGTATCAGCCTAAACATTGATGTCAGAGTCAATCATCTTTTGGAGCTCAATTCAGT 3352
Qy      1081  |||||ArgGlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGly 1100
Db      3353  |||||AGAGGAATTTGACACATTAATCTAGCCCTTATTAACCTTATGCACTGACAGGCGG 3412
Qy      1101  |||||SerProLysAlaLysGlnAlaLeuAsnMetLeuThrTyrArgAlaGlnGlnGlnGly 1120
Db      3413  |||||AGTCTTAAGCCAGAGAGACTTGAATATGCTGACTTGAGAGCCAGAACAGAGGTGAC 3472
Qy      1121  |||||MetGlnPheTyrValSerSerGlnSerLysLeuSerAspSerTyrGlnProArgSerLeu 1140
Db      3473  |||||ATGCAATTCGTGGGTGCATCAGAGTCCAAACTTTTGACTCCTGGCAGCCAGCTCCG 3532
Qy      1141  |||||AspIleGluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGlu 1160
Db      3533  |||||GATATTTAGAGTTGACGCCCTATGCACGCTCTCACTTCTTCAATTCAGACTTCTGAG 3592
Qy      1161  |||||GlyIleProIleMetArgTyrPheLeuSerArgGlnArgAsnSerLeuGlyPheAlaSer 1180
Db      3593  |||||GGAAATCCCAATTATGAGGTGGCTTAAGCAGGCAAGAAATAGCTTGAGGTGTTGACT 3652
Qy      1181  |||||ThrGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAlaAlaLeuMetAsnThr 1200
Db      3653  |||||ACTCAGAGATACCACTGTGGCTTTAAAGCTGTGTGATTTTGACAGCCCTAATGAAATA 3712
Qy      1201  |||||GluArgThrAsnIleGlnValThrValThrGlyProSerSerProSerPro----- 1217
Db      3713  |||||GAAAGGACAAATATCCAAAGTGAACGAGGCGGCTTACCAAGTCTCTTAAAGTTT 3772
Qy      1218  |||||-----LeuAlaValAlaGlnPro 1223
Db      3773  |||||CTGATTTGACACACACACCGCTTACTCTTCAGACACGAGAGCTTGCTGTGGTACAGCCA 3832
Qy      1224  |||||MetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValVal 1243

```

```

Db      3833  |||||ATGGCAATTAAATATTCCGCAAAATGTTTGGAATTTGCTAATTGTCACGCTCAATGTTGA 3892
Qy      1244  |||||TyrAsnValLysAlaSerGlySerSerArgArgArgSerIleGlnAsnGlnGluAla 1263
Db      3893  |||||TATATGTGAAGGCTTCTGGGTCTTCTGAAAGACGAAGATCTATCAAAATCAAGAGCC 3952
Qy      1264  |||||PheAspLeuAspValAlaValLysGluAsnLysAspAspLeuAsnHisValAspLeuAsn 1283
Db      3953  |||||TTTGATTTAGATGTTGCTGTAAGAAATAAAGATGATCTCAATCAATGATGATTTGAT 4012
Qy      1284  |||||ValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluValAsnLeu 1303
Db      4013  |||||GTGTGTACAAAGCTTTTCCGGCCCGGTGAGTGCATGCTTATATGGAAGTTAACTTA 4072
Qy      1304  |||||LeuSerGlyPheMetValProSerGluAlaIleSerLeuSerGluThrValLysVal 1323
Db      4073  |||||TTAAGTGCCTTATATGTCCTTCAGAACCAATTTCTGAGGAGACAGTGAAGAAAGTG 4132
Qy      1324  |||||GluTyrAspHisGlyLysLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCys 1343
Db      4133  |||||GAATATGATCAATGAAAACTCACTCATATTAGATTCGTAAATGAAGAAACCAAGTTTGT 4192
Qy      1344  |||||ValAsnIleProAlaValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSer 1363
Db      4193  |||||GTTAATATTCTCGCTGTGAGAAACTTTAAAGTTTCAATCCCAAGATGCTTCACTGTC 4252
Qy      1364  |||||IleValAspTyrTyrGlnProArgArgGlnAlaValArgSerTyrAsnSerGluValLys 1383
Db      4253  |||||ATAGTGATTAATAAGACCAAGAGACAGGCGGAGAAAGTTACAACTGAACTGAAG 4312
Qy      1384  |||||LeuSerSerCysAspLeuCysSerAspValGlnGlyCysArgProCysGluAspGlyAla 1403
Db      4313  |||||CTGTCTCTCTGAGACTTTGACGTGATGTCAGAGGCTCCGCTCTTGAGAGATGAGCT 4372
Qy      1404  |||||SerGlySerHisHisHisSerSerValIlePheIlePheCysPheLysLeuLeuTyrPhe 1423
Db      4373  |||||TCAGGCTCCCAATCAATCACTTCACTTCACTTATTTATTTCTGTTCAAGCTTCTGA 4432
Qy      1424  |||||MetGluLeuTyrPhe 1428
Db      4433  |||||ATGGAACCTTTGGCTG 4447

RESULT 9
AB079967
ID      AB079967 standard; cDNA, 5895 BP.
AC      AB079967;
DT      23-DEC-2002 (first entry)
DE      Human CD109 KI-H7 variant protein encoding cDNA.
KW      CD109; alpha2 macroglobulin; thioester; cerebroprotective; cardiant;
        immunosuppressive; haemostatic; anticoagulant; thrombolytic; human;
        cardiovascular; vasotropic; gene therapy; CD109 KI-H7; variant; gene; ss.
OS      Homo sapiens.
FH      Key
FT      Location/Qualifiers
FT      CDS
FT      /tag= a
FT      /product= "CD109 KI-H7 variant"
EN      WO200270696-A2.
PD      12-SEP-2002.
PF      07-MAR-2002; 2002MO-CA00292.
PR      07-MAR-2001; 2001US-273814P.
PA      (SCHU/) SCHU A.
        (SUTH/) SUTHERLAND R D.

```

XX Schuh A, Sutherland RD;  
PI WPI: 2002-713450/77.  
DR P-PSDB: ABB82168.  
XX New CD109 nucleic acids and polypeptides, useful in gene therapy,  
PT particularly for treating strokes, myocardial infarctions, thrombosis,  
PT thrombocytopenia, autoimmune diseases, or organ or bone marrow  
transplantation -  
XX Claim 1, Fig 2b, 156pp, English.  
XX The invention relates to isolated nucleic acid molecules encoding CD109  
CC polypeptides. These nucleic acid molecules include the human cDNA  
CC sequences comprising CD109 KI, CD109 KI-H7, CD109 K15 or their variants.  
CC CD109 is a novel member of the alpha2 macroglobulin (alpha2u)/C3, C4, C5  
CC family of thioester-containing proteins. The CD109 polypeptides can be  
CC expressed by standard recombinant methodology. The CD109 nucleic acid,  
CC CD109 polypeptide or its mimetic, or the anti-CD109 antibody is useful  
CC for treating or detecting a disease or disorder, e.g. conditions  
CC associated with endothelial activation, platelet activation, activation  
CC of the coagulation or fibrinolytic systems, activation of T lymphocytes  
CC and of the complement system, quantitative or qualitative abnormalities  
CC of platelet function, increased or impaired platelet aggregation and  
CC activation, increased or impaired activation of the coagulation and/or  
CC fibrinolytic systems, or impaired or increased immune activation. These  
CC are also useful for treating cardiovascular disorders, stroke, myocardial  
CC infarction, thrombosis, embolism, peripheral vascular disease,  
CC thrombocytopenia, thrombocytopenia, autoimmune diseases, organ  
CC transplantation, or bone marrow transplantation. The CD109 nucleic acids  
CC or their antisense nucleotide sequence are useful in gene therapy for  
CC treating these conditions. The present sequence represents the human  
CC CD109 KI-H7 variant cDNA sequence.

XX Sequence 5895 BP, 1708 A, 1174 C, 1241 G, 1772 T, 0 other;

## Alignment Scores:

Pred. No.: 0 Length: 5895  
Score: 7317.50 Matches: 1426  
Percent Similarity: 98.75% Conservative: 1  
Best Local Similarity: 98.69% Mismatches: 1  
Query Match: 99.58% Indels: 17  
DB: 24 Gaps: 1

US-10-020-095-4 (1-1428) x ABQ79967 (1-5895)

QY 1 MetGlnGlyProProLeuLeuThrAlaAlaHisLeuLeuGlyValGlyThrAlaAlaLeu 20  
DB 113 ATGCAGGGCCACCGCTCTGACCGCCGCCACCTCTCTGCTGTGACCGCGCGCTG 172  
QY 21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGly 40  
DB 173 GCCCGGCTCCCGGCTCGGCTTCTGTGTGACAGCCCGAGATCATGAGCCCGAGGA 232  
QY 41 AsnValThrIleGlyValGluLeuGluHisCysProSerGlnValThrValIleVal 60  
DB 233 AATGTGACTATTGGGGTGAAGCTTCTGSAACAATGCCCTTCAACAGTACTGTAAAGCG 292  
QY 61 GluLeuLeuGlyThrAlaSerAsnLeuThrValSerValLeuGluAlaGluGlyValPhe 80  
DB 293 GAGTGCTCAAGACAGCATCAAACTCACTGCTCTGCTCGAAGCAGAAAGAGTCTTT 352  
QY 81 GluIleGlySerPheLeuThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGlu 100  
DB 353 GAAAGAGGCTCTTTTAAACACTTACTTCCATCACTTCAAGACAGTGCAGATGAG 412  
QY 101 IleTyrGluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThr 120  
DB 413 ATTATAGCTACGTGTAAACGAGAGTACCCAGGATGAGATTTATTTCTTAATAGTACC 472  
QY 121 ArgLeuSerPheGluThrIleArgIleSerValPheIleGlnThrAspIleValLeuTyr 140

DB 473 CGCTTATCATTTGAGACCAAGAGATATCTGCTTCAATTCAAACAGACGCTTATAC 532  
QY 141 LysProGlyGlnGluValLysPheArgIleValThrIlePheSerAspPheProTyr 160  
DB 533 AAGCCAAAGCAAGAGTGAAGTTTCCGATTTTACACTCTTCTTCAGATTTTAAAGCTTAC 592  
QY 161 LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnIleTyrLeu 180  
DB 593 AAAACCTTTTAAACATTTCAATTAAGAGACCCCAATCAATTTGATCCAAAGTGGTTG 652  
QY 181 SerGlnIleSerAspLeuGlyValIleSerIleThrPheGlnLeuSerSerIleProIle 200  
DB 653 TCACCAACAAAGTATCTTGAAGTCAATTTCCAAAACCTTATGCTTCCATCCAAATA 712  
QY 201 LeuGlyAspTyrPheSerIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGln 220  
DB 713 CTGGTACTGGTCTTATTCAGTTCAAGTAAATGACCAAGCATTTTATCATCATTTTCAG 772  
QY 221 ValSerGluTyrValLeuProLysPheGluValThrIleGlnThrProLeuTyrCysSer 240  
DB 773 GTTTCAGATATGATATACCAAAATTTGAAGTACCTTGCAGACACCATTTATTTGTTCT 832  
QY 241 MetAsnSerIleHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProVal 260  
DB 833 ATGATTTCTAAGCATTTTAAATGGTACCATCAGCAAGATATGAGAAAGCCAGTG 892  
QY 261 LysGlyAspValThrLeuThrPheLeuProLeuSerPheTyrGlyLysLysAsnIle 280  
DB 893 AAAGAGACGTAAACGCTTACATTTTTCCTTTATCTTTGGGAAAGAAAGAAATATT 952  
QY 281 ThrIleThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluIleLys 300  
DB 953 ACAAAACATTTAAGATTAATGATGATCGCAAACTTCTTTTATATGATGAAGATGATAA 1012  
QY 301 AsnValMetAspSerSerAsnGlyLeuSerGluTyrIleAspLeuSerProGlyPro 320  
DB 1013 AATGTATGATTCCTCAAAATGAGCTTTCGAATACCTGTGATCTATCTTCCCTGAGCA 1072  
QY 321 ValGluIleLeuThrThrValThrGluSerValThrGlyIleSerArgAsnValSerThr 340  
DB 1073 GTAGAAATTTTAAACACAGTACAGATCACTTACAGTATTTTTCAGAAATGTAGACACT 1132  
QY 341 AsnValPhePheLysGlnHisAspTyrIleIleGluPhePheAspTyrThrThrValLeu 360  
DB 1133 AATGTTCTTCAAGCAACATGATTTACATCATTTGAGTTTTTGTATTTACTAGTCTTG 1192  
QY 361 LysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsnGlnLeu 380  
DB 1193 AAGCATCTCTCAACTTCAACAGCCACTGTGAAGTAACTGTCTGTATGCAACCAACTG 1252  
QY 381 ThrIleGluGluArgArgAsnAsnValIleThrValThrGlnArgAsnTyrThrGlu 400  
DB 1253 ACTCTTAAGAAAGAAATTAATGATGATCAATACATGACACAGAAACTTATCTGAG 1312  
QY 401 TyrTyrPheSerIleSerAsnSerGlyAsnGlnLysMetGluAlaValGlnLysIleAsnTyr 420  
DB 1313 TACTGAGCGGAGATCTAACAATGGAATCAGAAATGGAAGCTGTTCAGAAATTAATAT 1372  
QY 421 ThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGluAspSerSerGlu 440  
DB 1373 ACTGTCCCCCAAGTGGAACTTTTAAATGATTTCCCAATCTCTGGAGGATTTCCAGGAG 1432  
QY 441 LeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSerLeuPhe 460  
DB 1433 CTACAGTTGAAGGCTTATTTCTTGTGATGAAGATGCAATGAGTTCATAGTCTGTT 1492  
QY 461 LysSerProSerIleThrTyrIleGlnLeuLysThrArgAspGluAsnIleLysValGly 480  
DB 1493 AAGTCTCTAGTAAAGATATCAATCCAACTAAAAACAAGATGAAATATTAAGTGGGA 1552  
QY 481 SerProPheGluLeuValValSerGlyAsnLysArgLeuLysGluLeuSerTyrMetVal 500  
DB 1553 TCGCCTTTTGAAGTTGGTGTAGTGCAACAAAGATTTGAAGAGATTAAAGTATATGTA 1612

QY 501 ValSerArgGlyGlnLeuValAlaValGlyValGlnAsnSerThrMetPheSerLeuThr 520  
 DB 1613 GATCCAGGGGACAGTTGGTGGCTGAGGAAAACAAAATTCAAATGTTCTTTAACA 1672  
 QY 521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyrTrpIleGlnLysPheGly 540  
 DB 1673 CCGAATAATTTCTGGATCCCAAAAGCTTGTAATGGTATTAATTAAGATGATGGG 1732  
 QY 541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560  
 DB 1733 GAAATTAATAGATGATTTCTTAATAAATCTGTTACAGCTGTTTAAAAATTAAGATTAAG 1792  
 QY 561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580  
 DB 1793 CTATATTTGAGTAAGTGAAGCTGAACCATCTGAGAAGTCTCTTATGATCTCTGG 1852  
 QY 581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600  
 DB 1853 ACACAGCCCTGACTCCATGACTGGGATTTGAGCTGTGACAAAAGTGTGATCTGATGAAAT 1912  
 QY 601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeuGlnLeuTyrAsnThrGly 620  
 DB 1913 GCCTTAATGATTAATTAATGAAAAATGGTCCATGAGTTGGAACTTTATTAACACAGGA 1972  
 QY 621 TyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTrpVal 640  
 DB 1973 TATTTATTTAGCATGTTGATGAAATCTTTTGCAAGTCTTTCAGGAATGTGAGCTGGGTA 2032  
 QY 641 LeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyrAspAsnAlaGluTyr 660  
 DB 2033 TTGACAGATGACAACTCCAGAAAGATTTATTTGATGAGTGTATTATACATCAGAAATAT 2092  
 QY 661 AlaGluArgPheMetGluGlnLysGlnGlyHisIleValAspIleHisAspPheSerLeu 680  
 DB 2093 GCTGAGAGTTTATGAGGAAAAATGAGAACATTTGATGATTAATATACCTTTCTTTGG 2152  
 QY 681 GlySerSerProHisValArgLysHisPheProGluThrTrpIleTrpLeuAspThrAsn 700  
 DB 2153 GGTAGAGTCCACATGCTCCGAAAGCATTTTCCAGAGACTTGGATTGGCTAGACACCAAC 2212  
 QY 701 MetGlyTyrArgIleTyrGlnGlnPheGluValThrValProAspSerIleThrSerTrp 720  
 DB 2213 AAGGGTTCAGAGATTTTACCAAAATTTGAATTAAGTACGTACTGATCTATCACTCTGG 2272  
 QY 721 ValAlaThrGlyPheValIleSerGluAspLeuGlyLeuThrThrProVal 740  
 DB 2273 GTGGCTACTGTTTGTGATCTCTGAGAGCTGGGCTTGGACTTAACAACCTCCAGTG 2332  
 QY 741 GluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArg 760  
 DB 2333 GAGCTCCAGCTTCCACCACTTTTTCATTTTTCGAACTTCCCTACTCTGTATACAG 2392  
 QY 761 GlyGluGlnPheAlaLeuGlnIleThrIlePheAsnTyrLeuLysAspAlaThrGluVal 780  
 DB 2393 GGTGAAGAATTTGCTTTGAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2452  
 QY 781 LysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsn 800  
 DB 2453 AAGGTAATCATTTGAGAAAAAGTACAAATTTGATATTCTAATGACTTCAATGAAATAAAT 2512  
 QY 801 AlaThrGlyHisGlnGlnThrLeuLeuValProSerGluAspGlyValAlaThrValLeuPhe 820  
 DB 2513 GCCACAGCCACACAGACAGCCCTTCTGTTCCAGTGAAGATGGGCAACTGTTCTTTT 2572  
 QY 821 ProIleArgProThrHisLeuGlyGlnIleProIleThrValThrAlaLeuSerProThr 840  
 DB 2573 CCATCAGGCCAACACATCTGGGAGAAATTCATACAGTCAACAGCTCTTCAACCACT 2632  
 QY 841 AlaSerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIleGluLysSerTyr 860  
 DB 2633 GCTTGTGATGTGTCAACCAAGATGATTTTAAGTAAGCTGAAGGAATGAATAAATCATAT 2692

QY 861 SerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeu 880  
 DB 2693 TCACATATCATCTTTATGACTTGACTGACATAGGCTTACAGAGTACCTGAAACTTTG 2752  
 QY 881 SerPheSerPheProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIle 900  
 DB 2753 AGTTTCTCATTTTCTCCCTTAATACAGTGAAGTGGACAGTGAAGGATTCAGATCACTCAAT 2812  
 QY 901 GlyAspValLeuGluProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGly 920  
 DB 2813 GAGAGATGTTCTTGTCCTTCCATCAATGCTTACCTCATTTATTCGATGCTTATGGC 2872  
 QY 921 CysGlyGlnGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThr 940  
 DB 2873 TGTGGTGAACAGAACATGATTAATTTGCTCCAAATATTAATTAATTTGGATTAATGCACT 2932  
 QY 941 LysIleLysGlnLeuThrAspAsnLeuLysGlyValAlaLeuSerPheMetArgGlnGly 960  
 DB 2933 AAAAAAGAAACACATCAAGATTAATTTGAAGAAAAAGCTCTTTCAATTTATGAGCAAGCT 2992  
 QY 961 TyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyr 980  
 DB 2993 TACCAAGAGACCTTCTCTATCAGAGGAGAGATGCTCTTTCAGTGTCTTTGGAAATAT 3052  
 QY 981 AspProSerGlySerThrTrpLeuSerAlaPheValLeuArgCysPheLeuGluAlaAsp 1000  
 DB 3053 GACCTTCTGGAGGACCTTGGTGTGACGTTTGTATTAAGATGTTTCTTGAACCCAT 3112  
 QY 1001 ProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTrpLeuLysGlyHis 1020  
 DB 3113 CCTTCAATGATGATTTGATCAGATGCTGTACACAAACATACACTTGGCTTAAAGACAT 3172  
 QY 1021 GlnLysSerAsnGlyGluPheTrpAspProGlyArgValIleHisSerGluLeuGlnGly 1040  
 DB 3173 CAGAAATCCAGGAGGATTTTGGATCCAGGAAGATGATTCATAGTAGCTTCAAGT 3232  
 QY 1041 GlyAsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArg 1060  
 DB 3233 GGCATTAATAAGTCCAGTAACCTTACAGCTTAATTTGTAACCTTCTCTGGGATATAGA 3292  
 QY 1061 LysTyrGlnProAsnIleAspValGlnGlnSerIleHisPheLeuGlnSerGluPheSer 1080  
 DB 3293 AAGTATCAGCTTAACATGATGTCAGAGTCTATCACTTTTGGAGCTCAATTCAGT 3352  
 QY 1081 ArgGlyIleSerAspAsnTyrThrLeuAlaLeuIleTyrThrAlaLeuSerSerValGly 1100  
 DB 3353 AGAGAAATTTCCAGCAATTAATCTAGCCCTTAATTAATTAATTAATTAATTAATTAAT 3412  
 QY 1101 SerProLysAlaLysGluAlaLeuAsnMetLeuThrTrpArgAlaGluGlnGlnGlyGly 1120  
 DB 3413 AGTCTTAAGCAAGAAAGCTTTGAATATGCTGACTTGGAGAGCAAGAACAAAGTGGC 3472  
 QY 1121 MetGlnPheTrpValSerSerGluSerLysLeuSerAspSerTrpGlnProArgSerLeu 1140  
 DB 3473 ATGCAATCTGGGTCTCATCAGAGTCCAAACTTTTGCAGCTCTGGACAGCAGCTCCCTG 3532  
 QY 1141 AspIleGluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGlu 1160  
 DB 3533 GATATTTGAAGTTGACGCTTAATGACTGCTTCACTTCACTTCAATTTTCAAGCTTCTGG 3592  
 QY 1161 GlyIleProIleMetArgTrpLeuSerArgGlnArgAsnSerLeuGlyGlyPheAlaSer 1180  
 DB 3593 GGAAATCCCAATTAATAGGTGGCTAAGCAGGCAAAAGAAATAGCTTGGGTGGTTTGCATCT 3652  
 QY 1181 ThrGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAlaAlaLeuMetAsnThr 1200  
 DB 3653 ACTCAGATATCACATGCTGGCTTAAAGCTGTGTGATTTGACAGCTTATGATATACA 3712  
 QY 1201 GluArgThrAsnIleGlnValThrValThrGlyProSerSerProSerPro----- 1217  
 DB 3713 GAAAGACAAATATCAAGTGAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3772  
 QY 1218 -----LeuAlaValAlaGlnPro 1223

```

Db      3773 CTGATTGACACACACCGCTTACTCTTCAGACAGACGAGCTGTGTGTACAGCCA 3832
Qy      1224 MetAlaValaenileserAlaenGlyPhehailieCyseGlnLeuasnVala 1243
Db      3833 ATGGAGTTAAATATTCGCAAAATGGTTTGGATTTGCTAATTTGCACTCAATAGTTGTA 3892
Qy      1244 TyrAsnVallybAlaserGlySerSerzrgrArgrArgrSerileGlnAsnGlnuAla 1263
Db      3893 TATATATGGAAGGCTTCTGGGCTCTTCAGAGACGAATCTATCCAAATCAAGAGGCC 3952
Qy      1264 PheAspLeuAspValAlaVallyGlnAsnlyAspAspLeuAsnHisValAspLeuAsn 1283
Db      3953 TTTGATTTAGATGTTGCTGTAAGAAATAAAGATGATCTCAATCATGTGATTTTGAT 4012
Qy      1284 ValCyseThrSerPheSerGlyProGlyArgrSerGlyMetAlaLeuMetGlnValAsnLeu 1303
Db      4013 GTGTGTACAACTTTTCCGGCCCGGGTAGAGTGCGATGGCTTATATGAAATTAACTTA 4072
Qy      1304 LeuSerGlyPheMetValProSerGlnuAlaileSerLeuSerGluThrVallyLeuVal 1323
Db      4073 TTAAGTGCCTTTATGGTCCCTTCAGAACCAATTTCTCGACGACGACAGTGAAGAAAGTG 4132
Qy      1324 GluTyrAspHisGlyLyLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCys 1343
Db      4133 GAATATGATCATGGAACCTCAACCTCTAATTAGATCTGTAAATGAAGAACCCAGTTTGT 4192
Qy      1344 ValAsnleProAlaValArgrAsnPheValSerAsnThrGlnAspAlaSerValSer 1363
Db      4193 GTTAATATTCCTGCGTAGAGAACTTTAAAGTTTCAATACCCAGAGATCTTCAGTGTCC 4252
Qy      1364 IleValAspTyrTyrGluProArgrArgGlnAlaValArgrSerTyrAsnSerGlnuVallys 1383
Db      4253 ATAGTGGATTACTATGACCAAGAGAGACGGCGGTGAAGATTCAACTCGAAGTGAG 4312
Qy      1384 LeuSerSerCyAspLeuCySerAspValGlnGlyCyAspArgProCyGlnAspGlyAla 1403
Db      4313 CTGTCTCTCTGTGACCTTTGACATGATGTCCAGGGCTCCCTCTTGAGAGATGAGACT 4372
Qy      1404 SerGlySerHisHisSerSerValillePhelePheCyPheValLeuLeuTyrPhe 1423
Db      4373 TCAGGCTCCCATCATCAGCTTCAGTCAATTTTATTTCTGTTCAGCTTCGTACTTT 4432
Qy      1424 MetGluLeuTrpLeu 1428
Db      4433 ATGGAATTGGCTG 4447

```

RESULT 10  
AAA60199 standard; DNA; 4335 BP.

AC AAA60199;  
02-FEB-2001 (first entry)

Hydrophobic domain protein cDNA HP02837 isolated from HT-1080 cells.

Human, secreted protein; membrane protein; hydrophobic domain; proliferation control; differentiation induction; material transport; biophylaxis; signal receptor; ion channel; transporter; immunostimulant; immunosuppressant; haematopoiesis regulator; chemotactic; chemokinetic; haemostatic; thrombolytic; anti-inflammatory; tumor inhibition; autoimmune disease; Alzheimer's disease; Parkinson's disease; cancer; gene therapy; ss.

XX Homo sapiens.  
OS  
XX  
XX MO200029448-A2.  
XX  
XX 25-MAY-2000.  
XX  
XX 17-NOV-1999; 99WO-JP06412.  
XX

```

PR      17-NOV-1998; 98JP-0326255.  
PR      22-DEC-1998; 98JP-0364315.  
PR      16-MAR-1999; 99JP-0069811.  
PR      27-APR-1999; 99JP-0119299.  
PR      19-MAY-1999; 99JP-0138169.  
XX  
PA      (SAGA) SAGAMI CHEM RES CENT.  
PA      (PROT-) PROTEGENE INC.  
PI      Kato S, Kimura T;  
XX      WPI: 2000-387753/33.  
DR      P-PSDB; AAB12127.  
XX  
PT      Proteins comprising hydrophobic regions, such as secretory and membrane  
PT      proteins, useful in research and diagnostics and having various  
PT      activities e.g. immunomodulatory, antiinflammatory, chemokinetic,  
PT      hemostatic, thrombolytic -  
XX  
PS      Claim 3; Page 244-246; 410pp; English.  
XX  
CC      Secretory proteins play important roles in the proliferation control, the  
CC      differentiation induction, the material transport and the biophylaxis of  
CC      cells. Membrane proteins have important roles as signal receptors, ion  
CC      channels and transporters. The present sequence is the coding sequence  
CC      for a human protein which has at least one hydrophobic domain. The  
CC      protein encoded by the present sequence may be a secretory or a membrane  
CC      protein. The encoded protein may have cytokine and cell  
CC      proliferation/differentiation activity, immune stimulating or suppressing  
CC      activity, hematopoiesis activity, tissue growth activity,  
CC      activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC      and thrombolytic activity, anti-inflammatory activity and tumor  
CC      inhibition activity. The present sequence could therefore be used for  
CC      treatment of autoimmune disease, Alzheimer's disease, Parkinson's  
CC      disease, and cancer via gene therapy.  
XX  
SQ      Sequence 4335 BP; 1295 A; 838 C; 922 G; 1280 T; 0 other;  
XX  
Alignment Scores:  
Pred. No.: 0 Length: 4335  
Score: 7307.50 Matches: 1424  
Percent Similarity: 98.62% Conservative: 1  
Best Local Similarity: 98.55% Mismatches: 3  
Query Match: 99.45% Indels: 17  
DB: Gaps: 1  


```

US-10-020-095-4 (1-1428) x AAA60199 (1-4335)

```

Qy      1 MetGlnGlyProProLeuLeuThrAlaAlaHisLeuCyseValCyseThrAlaAlaLeu 20
Db      1 ATGCAAGGGCCCAACCGCTCGACCGCGCCCACTCTCTGCGTGTCAACCGCGGCTG 60
Qy      21 AlaValAlaProGlyProArgrPheLeuValThrAlaProGlyLleleArgrProGlyGly 40
Db      61 GCCGTGCTCCCGGCTCCGTTTGTGTGTACAGCCCAAGGATCATCAGAGCCCGAGGA 120
Qy      41 AsnValThrilleGlyValGlnLeuGlnHisCyseProSerGlnValThrVallyAla 60
Db      121 AATGTGCTATTTGGGGAGCTTCTGAAACATGCGCTTCACAGGTGATGTAAGGCG 180
Qy      61 GluLeuLeuysThrAlaSerAsnLeuThrValSerValLeuGlnuAlaGlnGlyValPhe 80
Db      181 GAGCTGTCAAGACAGATCAAACTACTGTCTGTCTGTGAGAGCAGAAGAGTCTTT 240
Qy      81 GluLyseGlySerPheLyThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGlu 100
Db      241 GAAAAAGCTCTTTTAAAGACACTTACTTCATCACTACTCTGAAACAGTGCAGATGAG 300
Qy      101 IleTyrGlnLeuArgrValThrGlyArgThrGlnAspGluLleuPheSerAsnSerThr 120
Db      301 ATTATAGCTACGTGTGAACCGAGCGTACCAGAGATGAGATTTTATTTCTTAATAGTACC 360
Qy      121 ArgLeuSerPheGluThrlyAsrGlyleSerValPheilleGlnThrAspLybAlaLeuTyr 140

```

```
Db 361 CGCTTATCATTTAGACCAAGAGAAATATGCTCTTCAATTCACAAACAAGGCTTATAC 420
Qy 141 LysProLysGlnGluValLysPheArgIleValThrLeuPheSerAspPheLysProLys 160
Db 422 AGCCCAAGCAAGAGAAAGTGAAGTTCCGATGTTTACACTCTTCACAGATTTTAAGCCTTAC 480
Qy 161 LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnIntProLys 180
Db 481 AAAACCTCTTAAACATTCCTCATTAAGGACCCCAAAATTCAAATTCACACAGTGGTTG 540
Qy 181 SerGlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProLys 200
Db 541 TCACAAACAAGAGATCTTGAGATCATTTCCAAAACCTTTTCAGCTATCTTCCCATCCAAA 600
Qy 201 LeuGlyAspTrpSerIleGlnValGlnValAsnAspGlnThrTyrTrpGlnSerPheGln 220
Db 601 CTGGTGACATGGTCTTATTCAGATTCAGATGACACAGACATATATCATCATTCATTCAG 660
Qy 221 ValSerGluTyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSer 240
Db 661 GTTTCAGAAATATGTATTTACAAAATTTGAAGTACCTTGACAGACACATTAATTTCTCT 720
Qy 241 MetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProVal 260
Db 721 ATGAATTCATGAAGATTTAAATGATGACATCACCGCAAGATATCATATGGGAAAGCCAGTG 780
Qy 261 LysGlyAspValThrLeuThrPheLeuProLeuSerPheTrpGlyLysLysAsnIle 280
Db 781 AAGAGAGAGCTTAACGCTTACATTTTACCTTATCCCTTTGGGAAAGAAAGAAATATAT 840
Qy 281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluMetLys 300
Db 841 ACAAATAACCTTAAAGATTAATGATGATGCAAACTTCTCTTTATATGATGAAGATGAAGA 900
Qy 301 AsnValMetAspSerSerAsnGlyLeuSerGlyTyrLeuAspLeuSerSerProGlyPro 320
Db 901 AATGTATGATGATCTTCAATGATGACCTTGAATACCTGATCTATCTTCCCTGACACA 960
Qy 321 ValGluIleLeuThrThrValThrGluSerValThrGlyIleSerTrpAsnValSerThr 340
Db 961 GTTAGAAATTTTAAACCAAGTACAGATGACAGATGATTAACAGTAAATGTAACCACT 1020
Qy 341 AsnValPhePheLysGlnHisAspTyrIleIleGluPhePheAspTyrThrThrValLeu 360
Db 1021 AATGTCTCTTCAAGACATGATTAATCATCATTCATGATTTTGAATTAATCTACTGCTTG 1080
Qy 361 LysProSerLeuAsnThrThrAlaThrValLysValThrArgAlaAspGlyAsnGlnLeu 380
Db 1081 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTGTGCTGATGCAACCAACTG 1140
Qy 381 ThrLeuGlnGluLysArgTrpAsnAsnValValIleThrValThrGlnArgAsnTyrThrGlu 400
Db 1141 ACTCTTGAAGAAAGAAAGAAATATATGATCATTAACAGTACACAGAGAAACTATATCTGAG 1200
Qy 401 TyrTrpSerGlySerAsnSerGlyAsnGlnLysMetGluAlaValGlnLysIleAsnTyr 420
Db 1201 TACTGAGGCGGATCTTACAGTGAATTCAGAAATTCGAAGCTGTTTCAGAAATTAATTAAT 1260
Qy 421 ThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGluAspSerSerGln 440
Db 1261 ACTGTCCTCCCAAGGTGAACCTTTTAAAGATTAATTCACAACTCTGAGAGATTCAGTGG 1320
Qy 441 LeuGlnLeuLysAlaLysPheLeuGlySerLysSerSerMetAlaValHisSerLeuPhe 460
Db 1321 CTACAGATTTAAAGGCTTAATTCCTTGGTAGTAAAGTAAAGATGACATGCTATATGCTGTT 1380
Qy 461 LysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLysValGly 480
Db 1381 AAGTCTCCATGATGACATACATCCAACTTAAACAAGAGATTAATTAAGGTGGGA 1440
Qy 481 SerProPheGluLeuValValSerGlyAsnLysArgLeuLysGluLeuSerTyrMetVal 500
Db 1441 TCGCCTTTTGAAGTGGTGTAGTGCAACAACGATTCGAAGAGTAAAGCTATATGCTGA 1500
Qy 501 ValSerArgGlyGlnLeuValAlaValGlyLysGlnAsnSerThrMetPheSerLeuThr 520
Db 1501 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAACAACAAATTCACAAATGTTCTCTTAA 1560
Qy 521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGluAspAspGly 540
Db 1561 CCAGAAATATTTGGACATCCCAAAAGCCTGTGTAAATTTGATTAATTAATGAATGATGGG 1620
Qy 541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560
Db 1621 GAAATTTAATGAAGATGTTCTAAATAATTCCTGTACGCTTGTAAAAATTAAGATTAAG 1680
Qy 561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580
Db 1681 CTATATTGAGTAAAGTGAAGTGAACCTTAACCATTTGAAAGTCTCTTATGGATCTCTG 1740
Qy 581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600
Db 1741 ACACAGCCTGACTCCATATGTTGGATTTGATGTGTTGACAAAAGTGAATCTGATGAT 1800
Qy 601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeuGlnLeuTyrAsnThrGly 620
Db 1801 GCCTTAATGATTAATTAACATGGAATAATGGTCCATGATGGAACTTTATTAACACAGGA 1860
Qy 621 TyrTyrLeuGluMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuThrVal 640
Db 1861 TATTAATTAAGGATGATTCATGAATTCCTTTGACGCTTTCAGGAATGGAGATCTGGGGA 1920
Qy 641 LeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyrAspAsnAlaGluTyr 660
Db 1921 TTGACAGATGCAAACTCACAGAGATTAATTAATGATGGTTTAAAGACATCAGAAATAT 1980
Qy 661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680
Db 1981 GCTGAGAGGTTTATGAGAGAAATAGAGACATATTAATTAATTAATCATGACTTTCTTGG 2040
Qy 681 GlySerSerProHisValArgLysHisPheProGluThrTrpIleTrpLeuAspThrAsn 700
Db 2041 GGTAGAGCTCCACAGTCCGAAAGCAATTTCCAGAGACTGTGATTTGGCTAACAACCAAC 2100
Qy 701 MetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIleThrSerTrp 720
Db 2101 ATGGTCTCCAGATTTTACAGAAATTTGAAGTACGTAACCTGATTCATCTCTTG 2160
Qy 721 ValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrProVal 740
Db 2161 GTGGCTACTGGTTTGTGATCTCTGAGGACCTGGGCTTTGGACTTAACAACATCTCCAGTG 2220
Qy 741 GluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArg 760
Db 2221 GAGCTCCAAAGCTTCCAAACCAATTTTCAATTTTGAATCTTCCCTACTCTGTATTCGA 2280
Qy 761 GlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrGluVal 780
Db 2281 GGTGAAGATTTGCTTTGGAATTAATCTATTAATTTGAAGAAGTCCACTAGGTT 2340
Qy 781 LysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsn 800
Db 2341 AAGGTATATCATTTGAAGAAAGTGAATTTGATTAATTAATTAATTAATTAATTAATTAAT 2400
Qy 801 AlaThrGlyHisGlnGlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPhe 820
Db 2401 GCCACAGGCCAACACAGACCCCTTCTGTGTTCCAGTGAAGATGGGCAACTGTTCTTTT 2460
Qy 821 ProIleArgProThrHisLeuGlyGluIleProIleThrValThrAlaLeuSerProThr 840
Db 2461 CCCATCAGGCCAACACATCTGGGAAATTTCTTATCAAGTACAGCTTTTACCCACT 2520
Qy 841 AlaSerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIleGluLysSerTyr 860
Db 2521 GCTTCTGATGCTATCACCAAGATGATTTTATGAAGGCTGAAGAGATTAAGAAATCATAT 2580
```

QY 861 SerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuIleThrLeu 880  
 Db 2581 TCACATTCATCTTATAGACTGACATAGCTACAGAGTACACCTGAAAACCTTGG 2640  
 QY 881 SerPheSerPheProPheAsnThrValThrGlySerGluArgValGlnIleThrAlaIle 900  
 Db 2641 AGTTCTCATTTCCCTCAATACAGTGAAGTGCAGTGAAGAGTTCAATCATCTGCAAT 2700  
 QY 901 GlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGly 920  
 Db 2701 GGAGATGCTTGGCTTCCATCATAGCTTAGCTCATTTGATTCGGATGCCCTTATGCG 2760  
 QY 921 CysGlyGlnGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThr 940  
 Db 2761 TGTGGTAAACAGAACATGATTAATTTGCTCAAAATATTTACATTTTGATTCGACT 2820  
 QY 941 LysIleIleValGlnLeuThrAspAsnLeuIleGlyValAlaLeuSerPheMetArgGlnGly 960  
 Db 2821 AAAAAGAAACAACTGACAGATTAATTTGAAGAAAAGCTCTTTCATTCAGGCAAGGT 2880  
 QY 961 TyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyr 980  
 Db 2881 TACCAAGAGAACTTCTATCAGAGGAGAGTGGCTCTTCAGTGTCTTTGGGAAATTA 2940  
 QY 981 AspProSerGlySerThrTyrLeuSerSerAlaPheValLeuArgCysPheLeuGluAlaAsp 1000  
 Db 2941 GACCTTCTGGGAGACCTTGGTGTGACCTTTTGTTCAGATGTTTCTTGAAGCCGAT 3000  
 QY 1001 ProTyrIleAspIleAspGlnAsnValLeuIleSarGlnTyrThrTyrLeuIleGlyHis 1020  
 Db 3001 CCTTACATAGATATGATCAGAAATGTGTACACAGAACATCACTTGGCTTAAAGGACAT 3060  
 QY 1021 GlnIleSerArgGlnGluPheThrAspProGlyArgValIleHisSerGluLeuGlnGly 1040  
 Db 3061 CAGAAATCAACGGGAAATTTGGATCCAGAGAAAGATTCATCAGTACCTTCAGAGT 3120  
 QY 1041 GlyAsnIleSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArg 1060  
 Db 3121 GGCATATAAAGTCCAGTAACTTACGCCCTTATNTGTACCTTCTCTCGGATATATGA 3180  
 QY 1061 LysTyrGlnProAsnIleAspValGlnGlnSerIleHisPheLeuGlnSerGluPheSer 1080  
 Db 3181 AAGATACAGCCATCATGATGTGCAAGAGTCACTTCTTGGAGTCTGAATTCAGT 3240  
 QY 1081 ArgGlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGly 1100  
 Db 3241 AGAGGAATTTCAAGACAAATTATACCTTACGCCCTTATACCTTATGCAATGTCAAGTGGG 3300  
 QY 1101 SerProIleValAlaValAlaLeuAsnMetLeuThrTyrPArgAlaGlnGlnGlyGly 1120  
 Db 3301 AGTCTTAAGAGGAGGAGGAGCTTGAATATGCTGACTGAGAGCAGAAACAAAGTGGC 3360  
 QY 1121 MetGlnPheTyrValSerSerGluSerIleLeuSerAspSerTyrGlnProArgSerLeu 1140  
 Db 3361 ATGCAATCTCGGGTGTCACTCAGAGTCCAACTTCTGACTCCTGGAGCCAGCTCCCTG 3420  
 QY 1141 AspIleGluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGlu 1160  
 Db 3421 GATATGAAAGTTGACGCTTATGCACTGCTTCACACTTCTTAAACATTTCAACATCTGAG 3480  
 QY 1161 GlyIleProIleMetArgTyrPheLeuSerArgGlnArgAsnSerLeuGlyGlyPheAlaSer 1180  
 Db 3481 GGAATCCCAATTAAGAGTGGCTAGACGAGCAAGAAATACCTTGGGTGGTTTTCATCT 3540  
 QY 1181 ThrGlnAspThrThrValAlaLeuIleValAlaLeuSerGluPheAlaLeuMetAsnThr 1200  
 Db 3541 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGACAGCCCTATAGATACA 3600  
 QY 1201 GluArgThrAsnIleGlnValThrValThrGlyProSerSerProSerPro----- 1217  
 Db 3601 GAAAGACAAATATTCAGAGTGAACGTCGAGGAGGCTTACCTCACAAGACCTGTAAAGTTT 3660

QY 1218 -----LeuAlaValAlaGlnPro 1223  
 Db 3661 CTGATTGACACACAAACCCCTTACTCTTCAGACAGACAGAGCTTGGTGTGATACAGCCA 3720  
 QY 1224 MetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValVal 1243  
 Db 3721 ACCGCACTTAATTTTCGCAAAATGTTTGGATTGCTATTTGTCAGCTCAATGTTGTA 3780  
 QY 1244 TyrAsnValIleAsnIleSerGlySerSerArgArgArgSerIleGlnAsnGlnGluAla 1263  
 Db 3781 TATATGTGAAGGCTTCTGGGCTTCTTACAGAGACGAAGATCTATCCAAATCAAGAGGCC 3840  
 QY 1264 PheAspLeuAspValAlaValIleGlyAsnIleAspAspLeuAsnHisValAspLeuAsn 1283  
 Db 3841 TTTGATTTCATGTTGCTGTAAAGAAATTAAGATCATCTCAATCATCTGTGGATTTCAT 3900  
 QY 1284 ValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluValAsnLeu 1303  
 Db 3901 GTGTGTCAAAGCTTTTGGGCCCCGGGTAGGAGTGGCATGCTCTTATGAGAAATTAACCTA 3960  
 QY 1304 LeuSerGlyPheMetValProSerGluAlaIleSerLeuSerGluThrValIleIleVal 1323  
 Db 3961 TTTAGTGGCTTTATGAGTGCCTTCAGAAACAAATTTCTGTGACGACAGACAGAAAGTGG 4020  
 QY 1324 GluTyrAspHisGlyIleLeuAsnLeuTyrLeuAspSerValAsnGlnThrGlnPheCys 1343  
 Db 4021 GAATATGATCATGGAACAACTCAACCTCTATTTAGATCTGTAAATGAACCCAGTTTGT 4080  
 QY 1344 ValAsnIleProAlaValAlaGlnAsnPheIleValSerAsnThrGlnAsnAlaSerValSer 1363  
 Db 4081 GTTAATATTCCTCTGTGAAACAACTTAAAGTTTCAATCCCAAGATGCTTCAAGTGTCC 4140  
 QY 1364 IleValAspTyrTyrGlnProArgArgGlnAlaValArgSerTyrAsnSerGluValIle 1383  
 Db 4141 ATTAGTGAATTACTATGAGCCCAAGAGACAGCCGCTGAGAACTTCAACTGTGAAGTGAAG 4200  
 QY 1384 LeuSerSerCysAspLeuCysSerAspValGlnGlyCysArgProCysGluAspGlyAla 1403  
 Db 4201 CTGTCTCTCTGTGACCTTTCAGATGATGTCAGAGGCTGCGTCTTGTGAGATGAGACT 4260  
 QY 1404 SerGlySerHisHisHisSerSerValIlePheIlePheCysPheIleLeuTyrPhe 1423  
 Db 4261 TCAGGCTCCCATATCATCACTTCAGTCATTTATTTCTGTTCAAGCTTCTGTACTTT 4320  
 QY 1424 MetGluLeuThrPhe 1428  
 Db 4321 ATGGAACCTTGGCTG 4335  
 Db  
 RESULT 11  
 ID AAA62010 standard; DNA; 4473 BP.  
 AC AAA62010;  
 XX  
 DT 02-FEB-2001 (first entry)  
 XX  
 DE Hydrophobic domain protein cDNA HP02837 isolated from HT-1080 cells.  
 XX  
 KW Human; secreted protein; membrane protein; hydrophobic domain;  
 KW proliferation control; differentiation induction; material transport;  
 KW biophysics; signal receptor; ion channel; transporter; immunostimulant;  
 KW immunosuppressant; haematopoiesis regulator; chemotactic; chemokinetic;  
 KW haemostatic; thrombolytic; anti-inflammatory; tumour inhibition;  
 KW autoimmune disease; Alzheimer's disease; Parkinson's disease; cancer;  
 KW gene therapy; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN MO200029448-A2.  
 XX  
 PD 25-MAY-2000.  
 XX  
 PF 17-NOV-1999; 99MO-JP06412.



XX 17-NOV-1998; 98JP-0326255.  
 PR 22-DEC-1998; 98JP-0364315.  
 PR 16-MAR-1999; 99JP-0069811.  
 PR 27-APR-1999; 99JP-0119299.  
 PR 19-MAY-1999; 99JP-0138169.  
 XX  
 PA (SAGA) SAGAMI CHEM RES. CENT.  
 PA (PROT-) PROTEGENE INC.  
 XX  
 PI Kato S, Kimura T;  
 DR WPI; 2000-387753/33.  
 DR P-PSDB; AAB12127.  
 XX  
 PT Proteins comprising hydrophobic regions, such as secretory and membrane  
 PT proteins, useful in research and diagnostics and having various  
 PT activities e.g. immunomodulatory, anti-inflammatory, chemokinetic,  
 hemostatic, thrombolytic -  
 XX  
 PS Claim 4; Page 253-261; 410pp; English.

CC Secretory proteins play important roles in the proliferation control, the  
 CC differentiation induction, the material transport and the biophysics of  
 CC cells. Membrane proteins have important roles as signal receptors, ion  
 CC channels and transporters. The present sequence is the coding sequence  
 CC for a human protein which has at least one hydrophobic domain. The  
 CC protein encoded by the present sequence may be a secretory or a membrane  
 CC protein. The encoded protein may have cytokine and cell  
 CC proliferation/differentiation activity, immune stimulating or suppressing  
 CC activity, haemopoiesis activity, tissue growth activity,  
 CC activating/inhibiting activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, anti-inflammatory activity and tumour  
 CC inhibition activity. The present sequence could therefore be used for  
 CC treatment of autoimmune disease, Alzheimer's disease, Parkinson's  
 CC disease, and cancer via gene therapy.

XX SQ Sequence 4473 BP; 1328 A; 867 C; 952 G; 1326 T; 0 other;

#### Alignment Scores:

Pred. No.: 0 Length: 4473  
 Score: 7307.50 Matches: 1424  
 Percent Similarity: 98.62% Conservative: 1  
 Best Local Similarity: 98.55% Mismatches: 3  
 Query Match: 99.45% Indels: 17  
 DB: 21 Gaps: 1

US-10-020-095-4 (1-1428) x AAA62010 (1-4473)

QY 1 MetGlnG1ProProleuLeuThra1Aa1Aa1sleuLeuCyvAlCyvThra1Aa1Aa1leu 20  
 DB 45 ATGCAAGGCGCCACCTCTCTGACCGCGCCACCTCTCTGCGTGGACCGCGCGGCG 104  
 QY 21 A1a1a1a1a1ProG1yProa1gPheLeuVal1Thra1AProG1y1le1le1a1a1a1ProG1y 40  
 DB 105 GCGTGGCTCCCGGCGCTCGGTTTCTGGTGAAGCGCCCAAGATCATCAAGCGCGGAG 164  
 QY 41 AenVal1Thr1le1y1Val1G1uLeuLeuG1u1le1CyPProSerG1nVal1ThrVal1y1a1a 60  
 DB 165 AATGTACTATTGGGGTGGAGCTTCTGGAACACTGCTGCTTCAACAGGTGACTGGAGG 224  
 QY 61 G1uLeuLeuYThra1Ase1AenLeuThra1Val1SerVal1leuG1u1aG1uG1yVal1Phe 80  
 DB 225 GACCTCTCAAGACACATCAACCTCACTGCTCTGCTCTGGAACAGAGAGCTTT 284  
 QY 81 G1uY1eG1ySerPhe1yThr1leuThr1leuProSer1e1uProLeu1a1a1a1a1a1a1a1a1 100  
 DB 285 GAAAAAGGCTCTTTAAGACACTTACTCTTCATCACTACCTCTGAACAGTGCAGATGAG 344  
 QY 101 l1e1y1G1uLeu1a1y1a1ThrG1y1a1g1Thr1a1a1a1G1u1leuPhe1Ser1a1a1a1 120  
 DB 345 ATTATGACTACGTGTAAACCGGACGTACCCAGATGAGATTATTTCTTAATAGTACC 404

QY 121 ArgLeuSerPheG1uThr1y1a1a1a1le1SerVal1Phe1leG1nThra1a1a1a1e1u1y 140  
 DB 405 CGCTTATATCTTGGAGCCAAAGAAATATCTGCTTCACTTCAAAACAGACAGGCTTATAC 464  
 QY 141 LysP1oLySG1nG1uVal1y1a1a1a1a1a1le1Val1ThrLeuPheSer1a1a1a1a1a1 160  
 DB 465 AAGCCAAAGCAAGAGTGAAGTTGCACTTGTACACTCTTCTCAGATTTTAAAGCTTAC 524  
 QY 161 LysThrSerLeu1a1a1leu1le1y1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1 180  
 DB 525 AAAACCTCTTAAACATCTCTAATAGAGACCCCAATCAATTTGATCCAAAGGTGTG 584  
 QY 181 SerG1nG1u1a1 200  
 DB 585 TCACAAACAAAGATCTTGGAGTCTTCTTCCAAACCTTTCAGCTATCTTCCATCCATAC 644  
 QY 201 LeuG1y1a1 220  
 DB 645 CTYGGTGAAGTGTCTTCAAGTTCAGTGAAGTGAACAGACATATATCAATCAATTCAG 704  
 QY 221 ValSerG1u1y1a1 240  
 DB 705 GTTTCAGAAATATGATATTCACAAATTTGAAGTGAAGTTCAGACACCATATATGTTCT 764  
 QY 241 MetA1a1 260  
 DB 765 ATGATTTCAACATTTAAATGATCCATCAAGCAAGTATACATATGCAAGCAAGT 824  
 QY 261 LysG1y1a1 280  
 DB 825 AAAGAGAGCTTAACCTTACATTTTACCTTATCTTTGGGGAAGAAAGAAATAT 884  
 QY 281 Thr1y1a1 300  
 DB 885 ACAAACATTTAAAT 944  
 QY 301 AenVal1MetA1a1 320  
 DB 945 AATGTAAATGATTTCTTCAATGATCTTCTGATATCTGATATCTTCTTCTTCTGACCA 1004  
 QY 321 ValG1u1le1uThr1Thra1Thr1u1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1 340  
 DB 1005 GTAGAAATTTTAAACAGAGTGAAGATCAAGTATCAAGTATTTCAAGAAATGAAGCACT 1064  
 QY 341 AenVal1Phe1y1a1 360  
 DB 1065 AATGTCTTCAACCAACATGATATATATATATATATATATATATATATATATATAT 1124  
 QY 361 LysProSer1e1u1a1 380  
 DB 1125 AAGCATCTCTAATCTTCAAGCCTGTAAGGTAAGTCTGCTGATGAGCAACCACTG 1184  
 QY 381 ThrLeuG1u1a1 400  
 DB 1185 ACTCTTGAAGAAAGAAAT 1244  
 QY 401 Tyr1TPSerG1ySer1a1 420  
 DB 1245 TACTGAGCGGATCTTCAAGTGAATATCAAGAAATGAAGCGTTCAAGAAATATATAT 1304  
 QY 421 ThrVal1ProG1nSerG1yThr1Phe1y1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1 440  
 DB 1305 ACTGCCCCCAAGTGAAGCTTTTAAATGATATATATATATATATATATATATATATAT 1364  
 QY 441 LeuG1nLeu1y1a1 460  
 DB 1365 CTACAGTTGAAGGCTTATTTCTTGTAGTAAAGATACATGAGCACTCATATGCTGTTT 1424  
 QY 461 LysSerProSer1y1Thr1y1a1 480  
 DB 1425 AAGTCTCTAGTAAACAT 1484  
 QY 481 SerProPheG1uLeuVal1a1 500

```
|||||
Db 1485 TCGCCTTTGAGTGGTGGTAGTGGCAAAACGATTGAAGGATTAAAGCTAATATGTA 1544
Qy ValSerArgGlyGlnLeuValAlaValGlySerGlnAsnSerThrMetPheSerLeuThr 520
Db 1545 GTATTCAGGGAGCACTTGTGGCTGTAGCAAAAACAAATTCACAAATTTCTCTTANCA 1604
Qy ProGluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGluAspArgI 540
Db 1605 CCAGAAATTTCTTGACTCCAAAGCCTGTGAATGTGTATATTAATGAAGATGATGG 1664
Qy GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560
Db 1665 GAAATTAAAGTGAATGTTCTAAAAATTCCTGTTCAGCTGTTTAAAAAATAAGATAAG 1724
Qy LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580
Db 1725 CTAATTCGAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 1784
Qy ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600
Db 1785 AACAGAGCTGACTCATATGTTGGGATGTGTACTGTGTACAAAGTAAGTAAGTAAGTAAG 1844
Qy AlaSerAsnAspIleThrMetGluAsnValValHisGluLeuGluLeuTyrAsnThrGly 620
Db 1845 GCCTCTAATGATATATTAATGAAATGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 1904
Qy TyrTyrLeuGluLysMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTrpVal 640
Db 1905 TATATTTAGGCAATGATATGATGAAATCTTTGCAAGCTTTCAGAAATGAGACTGGGTA 1964
Qy LeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyrAspAsnAlaGluTyr 660
Db 1965 TTGACAGATGCAAACTCAGCAAGGATTAATATGATGTTTATGACAAATGCAAGAAAT 2024
Qy AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680
Db 2025 GCTGAGAGTGTATGAGAAATGCAAGACATATGTAATATTCATGACCTTTCTTTTG 2084
Qy GlySerSerProHisValArgLysHisAspPheProGluThrTrpIleTrpLeuAspThrAsn 700
Db 2085 GGTACAGATCCACATGTCGAAAGCATTTTCCAGAGCTTGATGGCTGACCAAC 2144
Qy MetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIleThrSerTrp 720
Db 2145 ATGGTTCCAGGATTTACCAAGAAATTTGAACTGAACTGAACTGAACTGAACTGAACT 2204
Qy ValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrProVal 740
Db 2205 GTGGCTACTGGTTTGTGATCTCTGACGAGCCTGGGCTTGGACTAAACAACTACCTCCAGTG 2264
Qy GluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArg 760
Db 2265 GAGCTCCAAAGCTTCCAAACCAATTTTCATTTTGTGAATCTTCCCTACTCTGTATCCGA 2324
Qy GlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrGluVal 780
Db 2325 GGTGAAGAATTTGCTTTCGAATAATCAATATTCATATTTTGAAGAAGTCCAGAGGTT 2384
Qy LysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsn 800
Db 2385 AAGGTAATCATTTGAGAAAGTACAAATTTGATATTTCAATGACTTCAAGTAAGTAAGTAAG 2444
Qy AlaThrGlyHisGlnGlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPhe 820
Db 2445 GCCACAGGCCACACAGCAACCTTCTGTGTCCAGTGGAGATGGGCAACGTTCTTTT 2504
Qy ProIleArgProThrHisLeuGlyGluIleProIleThrValThrAlaLeuSerProThr 840
Db 2505 CCAATCAGGCCAACATCTGTGGAGAAATTTCTTATCAGATCAGAGCTTTTCCACCACT 2564
Qy AlaSerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIleGluLysSerTyr 860
|||||
Db 2565 GCTTCGATGCTATCACCAGATGATTTAGTAAGGCTGAGAGAAATGAAAAATCATAT 2624
Qy SerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeu 880
Db 2625 TCACAAATCATCTTATTAGACTTGACAGCAAAATAGGTACAGAGTACCCCGAAACCTTG 2684
Qy SerPheSerPheProAsnThrValThrGlySerGluArgValGlnIleThrAlaIle 900
Db 2685 AGTTTCTCATTTCTCTCTAATATACAGTACGAGGAGTGAAGAGTTGACATGCAAT 2744
Qy GluAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGly 920
Db 2745 GGAATGTTCTTGCTCTCTCATCAATAGGCTTACGCTTATGAGATGCTTATGGATGGC 2804
Qy CysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThr 940
Db 2805 TGTGTGAACAGAAACATGATTAATTTTGTCCCAAAATATTTACATTTTGGATTATCGACT 2864
Qy LysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPheMetArgGlnGly 960
Db 2865 AAAAAGAAACACAGTACAGATTAATTTGMAAGAAAGCTTTTCATTTATGAGGCAAGGT 2924
Qy TyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyr 980
Db 2925 TACCAGAGAACCTTCTATCAGAGGAAAGATGGCTTTTCACTGCTTTTGGAAATTA 2984
Qy AspProSerGlySerThrTrpLeuSerAlaPheValLeuArgCysPheLeuGluAlaAsp 1000
Db 2985 GACCTTCTTGAGACACTTGTGTGTGACGTTTGTCTTTTAAGATGTTTCCCTTGAAGCCGAT 3044
Qy ProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTrpLeuLysGlyHis 1020
Db 3045 CTTACATAGATATTTGATGCAATGTGTACAGAAATACATACCTTGCTTAAAGGACAT 3104
Qy GlnLysSerAsnGlyGluPheTrpAspProGlyArgValIleHisSerGluLeuGlnGly 1040
Db 3105 CAGAAATCCACAGGTGAATTTGGGATCCAGAAAGTGTATATGATGAGCTTCAAGGT 3164
Qy GluAsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArg 1060
Db 3165 GGCATTAATTAAGTCCAGTAACACTTACAGCCTTATTTCTTCTCCCTGGGATATAGA 3224
Qy LysTyrGlnProAsnIleAspValGlnGluSerIleHisPheLeuGluSerGluPheSer 1080
Db 3225 AAGTATCAGCCTAATGATGATGCAAGAGTATCATCAATTTTGGAGTGTGAATTCAGT 3284
Qy ArgGlyLysSerAspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGly 1100
Db 3285 AGAGGAATTTCAACAATTAATCTTACGCCCTTATTAATCATTTGTCATCAGTGGGG 3344
Qy SerProLysAlaLysGluAlaLeuAsnMetLeuThrTrpArgAlaGluGlnGlyGly 1120
Db 3345 AGTCTTAACAGAAAGAGAGCTTTGATATGCTGACTTGAGAGCAGAAACAGAGGTGGC 3404
Qy MetGlnPheTrpValSerSerGluSerLysLeuSerAspSerTrpGlnProArgSerLeu 1140
Db 3405 ATGCAATTTCTGGGATGATCAGAGTCCAAACTTTTGCAGCTTGAGCAGCCACTCCCTG 3464
Qy AspIleGluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGlu 1160
Db 3465 GATATTAAGAAATTCAGGCTATGCACTGCTTCACTCTTCAATTTTCAAGACTTTCGAG 3524
Qy GlyIleProIleMetArgTrpLeuSerArgGlnArgAsnSerLeuGlyIlePheAlaSer 1180
Db 3525 GGAATCCCAATTAATGAGTGGCTTAAGCAGGCAAGAAATAGCTTTGGGTGTTTGCATCT 3584
Qy ThrGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAlaAlaLeuMetAsnThr 1200
Db 3585 ACTCAGATTCACACTGTGGCTTTAAAGGCTCTGTGCAATTTGACAGCCCAATGAATACA 3644
Qy GluArgThrAsnIleGlnValThrValThrGlyProSerSerProSerPro----- 1217
Db 3645 GAAAGCAAAATATCCAAAGTACCGTACGCGGCTTACGCTCAACCAAGTCTCTTAAAGTTT 3704
```

```

QY 1218 -----Leu1aVal1aGlnPro 1223
Db 3705 CTGATTGACACACACACACCGCTTACTCTTACAGACGACGAGCTTCTGTGGTACAGCCA 3764
QY 1224 MetAlaValAsn1SerAlaanglypheGlyPheAla1leCysGlnLeuAenVal1 1243
Db 3765 ACCGAGATTAAATTTTCCCGCAAAVGGTTTGGATTGTCTATTTGTGAGCTCAATGTTGA 3824
QY 1244 TyrAsnVal1yalaSerGlySerSerArgArgArgSer1leGlnGnglnGln1a 1263
Db 3825 TATTAATGTAAGGCTTCTGGGCTTCTAGAAAGCAAGATCTATTCAAATCAAGAGCC 3884
QY 1264 PheAspLeuAspVal1aVal1yGlnAsn1yAspAspLeuAsn1sVal1AspLeuAen 1283
Db 3885 TTTGATTATGATCTCTCTGTAAAGAAATTAAGATGATCTCAATCATGTGCAATTTGAAT 3944
QY 1284 Val1CysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGlnVal1a 1303
Db 3945 GTGTGTACAAGCTTTTCGGGCGCGGTAGAGATGATGCTCTTATGGAAGTTAACCTA 4004
QY 1304 LeuSerGlyPheMetValProSerGlnAla1leSerLeuSerGlnThrVal1yGlyVal1 1323
Db 4005 TTAAGTGGCTTATAGATGCTTCAAGACAAATTTCTGAGCGAGCAAGTGAAGAAAGTG 4064
QY 1324 GluTyrAspHisGlyLysLeuAenLeuTyrLeuAspSerVal1aGlnThrGlnPheCys 1343
Db 4065 GATATGATCATGGAAGAACTCAACCTCTATTGATTTCTGTAAATTAACCAAGCTTTGT 4124
QY 1344 ValAsn1leProAlaVal1aArgAenPheVal1SerAsnThrGlnAsp1aSerVal1Ser 1363
Db 4125 GTTAATATTCCTGCTGTGAAGAACTTTAAAGTTTCAAAATACCAAGATGCTTCAAGTGC 4184
QY 1364 IleValAspTyrTyrGlnProArgArgGlnAlaVal1ArgSerTyrAsnSerGlnVal1y 1383
Db 4185 ATAGTGGATTACTATATAGCCAAAGACAGCGCGGTAGAAATTTTCAACTCTGAAGTGAAG 4244
QY 1384 LeuSerSerCysAspLeuCysSerAspVal1GlnGlyCysArgProCysGlnAspGlyAla 1403
Db 4245 CTGTCTCTCTGTACCTTTTGCAATGATGTCAGAGGCTGCGTCTGTGAGAGTGAAGCT 4304
QY 1404 SerGlySerHisHisSerSerVal1lePhe1lePheCysPheLysLeuLeuTyrPhe 1423
Db 4305 TCAGGCTCCCATCATCACTTTCAGTCATTTTATTTTCTGTTTCAAGCTTCTGACTTT 4364
QY 1424 MetGlnLeuTyrPhe 1428
Db 4365 ATGGAATTTGGCTG 4379

RESULT 12
AAD49436
ID AAD49436 standard; DNA; 4146 BP.
XX
AC AAD49436;
XX
DT 24-MAR-2003 (first entry)
XX
DE Human blood cell surface antigen, CD109 encoding DNA #1.
XX
KW Human; GPI-anchored TGF-beta1 binding protein; r150 protein; cancer;
KM glycosylphosphatidylinositol; transforming growth factor-beta1;
XX therapy; blood cell surface antigen; CD109; gene; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 1..4146
FT CDS.
FT /tag= a
FT /product= "Human CD109 protein"
FT /transl_except= (pos:2044..2046, aa:Xaa)
FT /note= "Xaa corresponds to Ser, Tyr; No start
FT and stop codon"
FT /partial

```

```

XX PN WO200285942-A2.
XX
XX 31-OCT-2002.
PD
XX
XX 24-APR-2002; 2002WO-CA00560.
XX PF
XX 24-APR-2001; 2001US-285713P.
XX PR
XX 14-FEB-2002; 2002US-356163P.
XX
XX (UWMC-) UNIV MCGILL.
XX PA
XX
XX Philip A, Tam B;
PI
XX
XX WPI: 2003-093100/08.
DR
DR P-PSDB; AAE32014.
XX
XX Novel transforming growth factor (TGF)-beta 1 binding reagent which
PT comprises r150 protein which acts as accessory receptor of TGF-beta,
PT useful for negatively modulating TGF-beta activity, and thus for
PT treating cancer
XX
XX Claim 7, Page 109-112; 127pp; English.
PS
XX
XX The invention relates to novel transforming growth factor (TGF)-beta1
CC binding reagent which comprises a glycosylphosphatidylinositol (GPI)-
CC anchored TGF-beta1 binding protein referred to as r150 which acts as
CC an accessory receptor of TGF-beta. The invention is used for negatively
CC modulating TGF-beta activity, and thus for treating conditions
CC characterised by overproduction of TGF-beta, such as cancer. Antisense
CC molecules of the invention are used for increasing TGF-beta availability
CC and increase graft success. The present sequence is human blood cell
CC surface antigen, CD109 encoding DNA. CD109 is a variant of r150 sequence.
XX
SQ Sequence 4146 BP; 1260 A; 776 C; 885 G; 1224 T; 1 other;
XX
Alignment Scores:
Pred. No.: 0 Length: 4146
Score: 7081.00 Matches: 1379
Percent Similarity: 99.86% Conservative: 1
Best Local Similarity: 99.78% Mismatches: 2
Query Match: 96.37% Indels: 0
DB: Gaps: 0

US-10-020-095-4 (1-1428) x AAD49436 (1-4146)
QY 22 ValAlaProGlyProArgPheLeuValThrAlaProGlyIleLeuArgProGlyGlyAsn 41
Db 1 GTGGCTCCCGGCGCTCGGTTTCTGTGACAGGCCAGGAGATCATAGGCCCGGAGGAAT 60
QY 42 ValThr1leGlyVal1GlnLeuLeuGlnHisCysProSerGlnVal1ThrVal1yAlaGlu 61
Db 61 GTGACTAATGGGGTGAAGCTTCTGTGAACATGCTCTTCAAGATGATGTAAGCGGAG 120
QY 62 LeuLeuLysThrAlaSerAsnLeuThrVal1SerVal1leGlnAlaGlnGlyVal1PheGlu 81
Db 121 CTGCTCAAGACGCAATCAACCTCACTGTCTGTGCTCGAAGACGAAGAGTCTTTGAA 180
QY 82 LysGlySerPheLysThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGlu1le 101
Db 181 AAAGCTCTTTTAAACACCTTACTTCTTCAATCAGTCTGGAAGAGTGAAGTGAAGT 240
QY 102 TyrGlnLeuArgVal1ThrGlyArgThrGlnAspGlu1leLeuPheSerAsnSerThrArg 121
Db 241 TATGAGCTACGTGTAAACCGGACGTACCCAGATGATGATTTTCTTAATATGATCCCGC 300
QY 122 LeuSerPheGlnThrLysArg1leSerVal1Phe1leGlnThrAspLysAlaLeuTyrLys 141
Db 301 TTATCATTTTGAACCAAGAAATATCTGTCTTCAATTAACAGACAGAGCCCTTATACAG 360
QY 142 ProLysGlnGlnVal1yPheArg1leVal1ThrLeuPheSerAspPheLysProTyrLys 161
Db 361 CCAAGCAAGAAAGTGAAGTTTTCGCAATTTGTTCACCTCTTCAAGATTTTAAGCTTACAA 420

```

162 ThrsleuAenlleleuileysaPProlySerAenlleuileglnlntPrleuSer 181  
421 ACCTCTTAACATCTTCATTAGAGACCCCAATTAATTGATCCAAACAGTGGTGC 480  
182 GlnGlnSerAapleuGlyValileSerlyThrPheGlnleuSerSerHieProleu 201  
481 CAACAAGTGAATCTTGAGTCATTTCCAAAATTTTCAGTATCTTCCATCCAAATCTT 540  
202 GlyAapTPSerileGlnValGlnValAsnAapGlnThrTyTyGlnSerPheGlnVal 221  
541 GGTGCTGCTCATTCAGTTCAGTGAATGACAGACATACATCAATTCATTCAGGTT 600  
222 SerGlnTyValleuProlyPheGlnValThrleuGlnThrProleuTyCyseSerMet 241  
601 TCAGAAATATGATTAACAAAATTTGAGTGAATCTTGACAGACACATTAATTTGTTATG 660  
242 AenSerlySHleuAenGlyThrileThrAlaTyTyThrTyGlyLysProValLys 261  
661 AATTTAAGCATTTTAAATGTATCCATTCACGGCAAAAGTATACATATGGGACCCAGTGAA 720  
262 GlyAapValThrleuThrPheleuProleuSerPheTPGlyLysLysAenlleThr 281  
721 GAGACGTAACGCTTACATTTTACCTTTATCTTTGGGAAAGAAAGAAAATATTAACA 780  
282 LysThrPheLysileAenGlySerAlaAenPheSerPheAenAapGlnGlnMetLysAsn 301  
781 AAAACATTAATAGATTAATAGATTCGCAAACTTCTTTAATAGTAAGATGAAAAAT 840  
302 ValMetAapSerSerAenGlyLysSerGlnTyLysAapleuSerSerProGlyProVal 321  
841 GTATATGATCTTCAATAGCATTTCTGAAATACCTGATCATCTTCCCTGGACCAAGTA 900  
322 GlnleuLeuThrThrValThrGlnSerValThrGlyileSerArgAenValSerThrAsn 341  
901 GAAATTTTAACACAGTACAGAAATCAGTACAGTATTTCAAGAAAGTAAGACATTAAT 960  
342 ValPhePheLysGlnHieAapTyTyileleGlnPhePheApyTyThrThrValleuLys 361  
961 GTGTTCTTCAAGCAACATGATTAATCATGATTTTGTATTAATCATCATCTCTTGAAG 1020  
362 ProSerleuAenPheThrAlaThrValLysValThrArgAlaAapGlyAenGlnleuThr 381  
1021 CCATCTCTCAACTTCAACAGCACCTGTGAAGGTAACTCTGCTGATGGCAACCACTGACT 1080  
382 LeuGlnLysArgAenAenValValileThrValThrGlnArgAenTyThrGlnTyTy 401  
1081 CTTGAAGAAAGAAAGAAATATGATGTCATTAACAGTACACAGAAACATTAATGAGTAC 1140  
402 TrpSerGlySerAenSerGlyAenGlnLysMetGlnAlaValGlnLysileAenTyTyThr 421  
1141 TCGAGCGGATCTTAACAGTGAATATCAAGAAATGAGAGCTGTTCAAGAAATTAATTAAT 1200  
422 ValProGlnSerGlyThrPheLysileGlnPheProileuGlnAapSerSerGlnleu 441  
1201 GTCCCCCAAAAGTGAATCTTTAAGTGAATTTCCCAATCTCGAGAGATTCAGAGAGCTA 1260  
442 GlnleuLysAlaTyTyPheLysGlySerlySerSerMetAlaValHieSerleuPheLys 461  
1261 CAGTTGAAGGCTTATTTCTTGTGATGAATAAGTACATGAGTTCATTAATGCTTTTAAG 1320  
462 SerProSerlyThrTyTyileGlnleuLysThrArgAapGlnAenlleLysValGlySer 481  
1321 TCTCTAGTAAAGACATCACTCACTCACTAAAACCAAGATGAAAAATTAATTAAGTGGATCG 1380  
482 ProPheGlnleuValValSerGlyAenLysArgleuTyGlnleuSerTyMetValVal 501  
1381 CCTTTTGAAGTGGTGTAGTGGCAACAAAGATTAAGAGATTAAAGTATATATGATGTA 1440  
502 SerArgGlyGlnleuValAlaValGlyLysGlnAenSerThrMetPheSerleuThrPro 521  
1441 TCCAGGGGACAGTGTGGCTGTGAGAAAAACAAAATTCACAAATGTTCTCTTTAAACACA 1500

522 GlnAenSerTrpThrProlyAlaCyseValileValTyTyTyTyileGlnAapGlyGln 541  
1501 GAAATTTCTTGACCTCCAAAAGCTGTGTAATGTGATTAATTAATGAATGATGGGAA 1560  
542 IleIleSerAapValleuLysileProValGlnleuValPheLysAenLysileLysleu 561  
1561 ATTATAGTATGATTTCTAAAATTTCTGTCACCTGTGTTTAAAAATGAATGAATGCTA 1620  
562 TyTrpSerlyValLysAlaGlnProSerGlnLysValSerleuArgileSerValThr 581  
1621 TATTTGAGTAAAGTGAAGCTGAAACCATCTGABAAAGTCTCTTTAGATCTCTTGACA 1680  
582 GlnProAapSerileValGlyileValAlaValAapLysSerValAenleuMetAenAla 601  
1681 CAGCTTACCTCATTAGTGGATTTGATGCTGTGACAAAAGTGAATCGAATGAAATGCC 1740  
602 SerAenAapIleThrMetGlnAenValValHieGlnleuGlnleuTyTyAenThrGlyTy 621  
1741 TCTAATGATATTAACAATGAAATGTGTCCATGAGTGGAACTTTATTAACACAGATAT 1800  
622 TyTrpGlnLysPheMetAenSerPheAlaValPheGlnGlnLysGlyLysleuThrValleu 641  
1801 TATTTAGGATGCTTCATGAATCTTTTGACGCTTTTCAGGAATGTGACCTGTGGTATTC 1860  
642 ThrAapAlaAenleuThrlyAapTyTyileAapGlyValTyAapAenAlaGlnTyTyAla 661  
1861 ACAGATCAAACTCCACGAAGATTAATGATGATGATTTATGATGATGATGATGATGCT 1920  
662 GlnArgPheMetGlnLysGlnGlnLysHieileValAapIleHieAapPheSerleuGly 681  
1921 GAGAGTTTATGAGAGAAATGAAGACATATGTGATTAATGATGATGATGATGATGCT 1980  
682 SerSerProHieValArgLysHiePheProGlnThrTrpIleThrleuAapThrAenMet 701  
1981 AGCAGTCCACATGCTCCGAAGACATTTCCAGACATTTGATGCTGATGACACCAACATG 2040  
702 GlyTyArgileTyGlnleuPheGlnValThrValProAapSerileThrSerTrpVal 721  
2041 GGTTCAGATTAATTAACGAATTTGAAGTATGATGATGATGATGATGATGATGATGCT 2100  
722 AlaThrGlyPheValileSerGlnAapleuGlyLysleuThrThrThrProValGln 741  
2101 GCTATCGTGTGTGATCTCTGAGACCTGGGCTTGGATTAACATCACTCACTCACTGAG 2160  
742 LeuGlnAlaPheGlnProPhePheIlePheLysAenleuProTyTySerValileArgGly 761  
2161 CTCAGACCTTCCAAACATTTTTCATTTTTCATTTTTCATTTTTCATTTTTCATTTTTC 2220  
762 GlnGlnPheAlaGlnGlnIleThrIlePheAenTyTyLysAapAlaThrGlnValLys 781  
2221 GAAATTTGCTTGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2280  
782 ValIleIleGlnLysSerAapLysPheAapIleleuMetThrSerSerGlnIleAenAla 801  
2281 GTATATCTTGAAGAAAGTGAACAAATTTGATTAATGATGATGATGATGATGATGATG 2340  
802 ThrGlyHieGlnGlnThrleuLysValProSerGlnAapGlyAlaThrValleuPhePro 821  
2341 ACAGGCCACACAGACACCTCTGTGTTCCAGTGAAGATGGGCACTGCTTTTTCCTCC 2400  
822 IleArgProThrIleleuGlnGlnIleProIleThrValThrAlaAenSerProThrAla 841  
2401 ATCAGGCCAACATCTGGGAAATTTCCATCACTCACTCACTCACTCACTCACTCACT 2460  
842 SerAapAlaValThrGlnMetileuValLysAlaGlnGlyLysleuLysSerTyTySer 861  
2461 TCTGATGCTGTCAACCAAGATTTTAAGTGAAGGCTGAAGGATTAAGAAATCATATTTCA 2520  
862 GlnSerIleleuLysAapleuThrAapAenArgleuGlnSerThrleuLysThrleuSer 881  
2521 CAATTCATCTTATGATCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2580  
882 PheSerPheProAenThrValThrGlySerGlnArgValGlnIleThrAlaIleGly 901

Db 2581 TTCATCTTCCTCTCATATACAGTACGTGACGAGAAAGATTCAATCTCAATTTGAA 2640  
 Qy AspValIleuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGlyCys 921  
 Db 2641 GATGTTCTTGTCCTTCCTCATCAATGAGCTTACCTCATTAATTCGATAGCTTAATGAGCT 2700  
 Qy 922 GlyGluGlnAsnMetIleAsnPhenAlaProAsnIleTyrIleuAspTyrLeuThrIlys 941  
 Db 2701 GGTGAACAGACATGATTAATTTTGTCTCAAAATATTACATTTTGGATTATCTGACTAAA 2760  
 Qy 942 LysIleGlnIleuThrAspAsnLeuGlyIleuValAlaLeuSerPheMetArgIleGlyTyr 961  
 Db 2761 AAGAAACCACTGACACATTAATTTGAAAAGAAAGCTTTTCATTATGAGGCAAGGTTAC 2820  
 Qy 962 GlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyrAsp 981  
 Db 2821 CAGAGAACTTCTCTATCAGAGGAGAGATGGCTTTTCAGTGTCTTTGGAAATTAATGAC 2880  
 Qy 982 ProSerGlySerThrTyrLeuSerAlaPheValIleuArgCysPheLeuGluAlaAspPro 1001  
 Db 2881 CTTTCTGGAGGACCTTGTTGTGACGTTTGTGTTTAAAGATGTTTCTTGAGCCGATCT 2940  
 Qy 1002 TyrIleAspIleAspGlnAsnValIleuHISArgThrTyrThrTyrLeuIysGlyHisGln 1021  
 Db 2941 TACATGATTAATGATCAGATGTTGTTACACAGAACATACACTGGCTTAAAGACATCAG 3000  
 Qy 1022 LysSerAsnGlyGluPheTyrAspProGlyArgValIleHisSerGluLeuGlnGly 1041  
 Db 3001 AATATCCAGCGTGAATTTTGGATCCAGAGAGTATTCATATGATGTTCAAGGTGAGC 3060  
 Qy 1042 AsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArgLys 1061  
 Db 3061 AATTAAGTCCAGTACACTTACAGCTTATATTGTAATCTTCTCTGGGATTAAGAAAG 3120  
 Qy 1062 TyrGlnProAsnIleAspValGlnGlnSerIleHisPheLeuGlnSerGluPheSerArg 1081  
 Db 3121 TATCAGCCTTAACATTGATGTCAGAGAGTCAATCATTTTTGGAGCTGAATTCAGTAGA 3180  
 Qy 1082 GlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGlySer 1101  
 Db 3181 GGAATTTACAGACATTAATCTAGCCCTTAATACTTAATGCAATGTCATCAGGGGAGAT 3240  
 Qy 1102 ProLysAlaLysGluAlaLeuAsnMetLeuThrTyrArgAlaGluGlnGlyGlyMet 1121  
 Db 3241 CCTAAGCGAAGAACCTTGAATATGTCGACTTGAAGAGCAGACAAAGAGTGCATG 3300  
 Qy 1122 GlnPheTyrValSerSerGlnSerIleuSerAspSerTyrGlnProArgSerLeuAsp 1141  
 Db 3301 CAATCTGGGTGCATCAGAGTCCAAACTTTCTGACTCTGGCAGCCAGCTCCCTGGAT 3360  
 Qy 1142 IleGluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGly 1161  
 Db 3361 ATTTGAAGTTGACGCTTATGCTACGCTCTCACTTCTTCAATTTTCAGACTTCTGAGGGA 3420  
 Qy 1162 IleProIleMetArgTyrLeuSerArgGlnArgAsnSerLeuGlyGlyPheAlaSerThr 1181  
 Db 3421 ATCCCAATTAATGAGTGGCTTACAGGCAAGAAATAGCTGGGTGCTTTGCAATTAAT 3480  
 Qy 1182 GlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAlaAlaLeuMetAsnThrGln 1201  
 Db 3481 CAGGATACCACTGTGCTTTAAAGCTGTCTGTGATTTTGCAGCCCTTAATGAATACAGA 3540  
 Qy 1202 ArgThrAsnIleGlnValThrValThrGlyProSerSerProSerProLeuAlaVal 1221  
 Db 3541 AGGACAAATATCCAACTGACCGTGAAGGGGCTTACGCTCAACAAATCTCTTGTGTGGTA 3600  
 Qy 1222 GlnProMetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsn 1241  
 Db 3601 CAGCCAAAGCAATTAATTTTCGCAAAATGGTTGGATTTGCTATTTGTGACGCTCAAT 3660  
 Qy 1242 ValValTyrAsnValLysAlaSerGlySerSerArgArgArgSerIleGlnAsnGln 1261

Db 3661 GTTGTATTAATATGAGGCTTGGGTCTTCTTACAGACAGAAATCTATCCAAATCA 3720  
 Qy 1262 GluAlaPheAspLeuAspValAlaValIysGluAsnLysAspAspLeuAsnHisValAsp 1281  
 Db 3721 GAAGCTTTGATTTTGAATGCTTGTCTTAATAAGAAATTAAGATGATCTCAATCATGTGAT 3780  
 Qy 1282 LeuAsnValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluVal 1301  
 Db 3781 TTGAATGATGTACAAAGCTTTTCCGGCCCGGGTAGAGATGGCATGGCTTATGGAAGTT 3840  
 Qy 1302 AsnLeuLeuSerGlyPheMetValProSerGluAlaIleSerLeuSerGluThrValIys 1321  
 Db 3841 AACCTATTAAAGTGGTTTATGCTGCTTCAGAAAGCAATTTCTTGAAGGAGACAGTGAAG 3900  
 Qy 1322 LysValGluTyrAspHisGlyLysLeuAsnLeuTyrLeuAspSerValAsnGluThrGln 1341  
 Db 3901 AAGTGGAAATGATCATGAGAAACTCAACTCTTAATTTAGATTTCTGTAAAGAAACCAAG 3960  
 Qy 1342 PheCysValAsnIleProAlaValArgAsnPhenValSerAsnThrGlnAspAlaSer 1361  
 Db 3961 TTTGTGTTAATATTCCTGCTGTGAGAACTTAAAGTTCAAAATACCAAGATGCTTCA 4020  
 Qy 1362 ValSerIleValAspTyrTyrGluProArgArgGlnAlaValArgSerTyrAsnSerGlu 1381  
 Db 4021 GTGTCCATAGTGGATTAATGATGAGCCAGAGACAGCGGTGAGAGATTACCAACTGTGA 4080  
 Qy 1382 ValLysLeuSerSerCysAspLeuCysSerAspValGlnGlyCysArgProCysGluAsp 1401  
 Db 4081 GTGAAGCTGTCTCTGAGCCTTTCAGTATGTCAGTATGTCAGAGGCTGCCGCTTGTGAGAT 4140  
 Qy 1402 GlyAla 1403  
 Db 4141 GGAAGCT 4146  
 RESULT 13  
 ID AAD49437  
 AAD49437 strand; DNA; 4197 BP.  
 AC AAD49437;  
 DT 24-MAR-2003 (first entry)  
 XX Human r150 DNA #2.  
 DE Human r150 DNA #2.  
 KW Human; GPI-anchored TGF-beta1 binding protein; r150 protein; cancer;  
 KW glycosylphosphatidylinositol; transforming growth factor-beta1;  
 KW therapy; gene; ds.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT CDS 1..4197  
 FT /\*tag= a  
 FT /product= "Human r150 protein #3"  
 FT /transl\_except= (pos:2044..2046, aa:Xaa)  
 FT /note= "Xaa corresponds to Ser, Tyr; No start  
 FT and stop codon"  
 FT /partial  
 PN WO200285942-A2.  
 PD 31-OCT-2002.  
 XX 24-APR-2002; 2002WO-CA00560.  
 PF 24-APR-2001; 2001US-285713P.  
 PR 14-FEB-2002; 2002US-356163P.  
 XX (UWMC-) UNIV MCGILL.  
 PA Phillip A, Tam B;  
 PI WPI; 2003-093100/08.  
 DR

DR P-PSDB: AAB32015.

XX Novel transforming growth factor (TGF)-beta 1 binding reagent which  
PT comprises r150 protein which acts as accessory receptor of TGF-beta,  
PT useful for negatively modulating TGF-beta activity, and thus for  
PT treating cancer -

XX Claim 7, Page 118-120; 127pp; English.

XX The invention relates to novel transforming growth factor (TGF)-beta1  
CC binding reagent which comprises a glycosylphosphatidylinositol (GPI)-  
CC anchored TGF-beta1 binding protein referred to as r150 which acts as  
CC an accessory receptor of TGF-beta. The invention is used for negatively  
CC modulating TGF-beta activity, and thus for treating conditions  
CC characterised by overproduction of TGF-beta, such as cancer. Antisense  
CC molecules of the invention are used for increasing TGF-beta availability  
CC and increase graft success. The present sequence is human r150 protein  
CC encoding DNA.

XX Sequence 4197 BP; 1276 A; 788 C; 894 G; 1238 T; 1 other;

Alignment Scores:

Pred. No.:	0	Length:	4197
Score:	7068.50	Matches:	1380
Percent Similarity:	98.71%	Conservative:	1
Best Local Similarity:	98.64%	Mismatches:	17
Query Match:	96.20%	Indels:	1
DB:	25	Gaps:	1

US-10-020-095-4 (1-1428) x AAD49437 (1-4197)

QY 22 ValAlaProGlyIProArgPheLeuValThrAlaProGlyIleIleArgProGlyIValAsn 41  
Db 1 GTGGCTCCCGGGCCCTCGGTTTCTGGTGACAGCCCGAGGATCATGAGCCCGAGGAAT 60  
QY 42 ValThrIleGlyValGluLeuLeuGluHisCysProSerGlnValThrValIleAlaGlu 61  
Db 61 GTGACTATTGGGGTGAGACCTTCTGGACACAGCCCTTCACAGGAGTGTAAAGCGGAG 120  
QY 62 LeuLeuValThrAlaSerAsnLeuThrValSerValIleGluIleAlaGluValPheGlu 81  
Db 121 CTGGCCAAAGACGACATCAACCTCATCTCTGCTGCGAAGCAAGAGGAGCTTTGAA 180  
QY 82 LysGlySerPheLysThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGluIle 101  
Db 181 AAAAGCTCTTTAAGACACTTACTCTTCATCATCACTCTGAAAGAGGAGATGAGATT 240  
QY 102 TyrGluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThrArg 121  
Db 241 TATGAGCTACGCTGTAAACCGAGACGTACCCAGATGAGATTATTCTTAATAGTACCGC 300  
QY 122 LeuSerPheGluThrThrValArgIleSerValPheIleGlnThrAspValIleuTyrLys 141  
Db 301 TTATCATTTTGAACCAAGAAATATCTGTCTTCAATCAACAGCAAGGCTTATACAG 360  
QY 142 ProLysGlnGluValIlePheArgIleValThrLeuPheSerAspPheLysProTyrLys 161  
Db 361 CCAAGCAAGAGTGAAGTTTCGCACTTGTACACTCTTCAAGATTTTAACCTTTACAA 420  
QY 162 ThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGlnTyrLeuSer 181  
Db 421 ACCCTTTAAACATCTCATTAAGAGACCCCAATCAATTTGATCCAAACAGTGTGCA 480  
QY 182 GlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIleLeu 201  
Db 481 CAACAAAGTGAATCTTGAGTCAATTCCTTCCAACTTTCCTCCATCCAAATACTT 540  
QY 202 GlnAspTrpSerIleGlnValGlnValAsnAspGlnThrTyrArgIleSerPheGlnVal 221  
Db 541 GGTAAGTGGTCTATTCAAGTTCAAGTGAATGACAGACATATTTCAATCATTTCAAGTT 600  
QY 222 SerGlnTyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSerMet 241  
|||

Db 601 TCAGAAATATGTATTACCAAAATTTGAAGTACTTTGCAGACACCATTAATTTGTTATG 660  
QY 242 AsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProValLys 261  
Db 661 AATTCTAAGATTTAAATGTTATCATCAGCGCAAGATATCATTTGGAGCCAGTGAA 720  
QY 262 GlyAspValThrLeuThrPheLeuProLeuSerPheTrpGlyLysLysAsnIleThr 281  
Db 721 GGAGACGTAAAGCTTCAATTTTACCTTTATCCCTTTGGGAAAGAAAGAAATATTACA 780  
QY 282 LysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluIleuMetLysAsn 301  
Db 781 AAAACATTTAAGATTAATGATGATTCGAAACTTCTTTTATATGATGAAGAGTAAAT 840  
QY 302 ValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSerSerProGlyProVal 321  
Db 841 GTAAATGATTTCTCAATGATGATCTTGTGAATACCTGATGATTAATCTCCCTGAGCAGTA 900  
QY 322 GluIleLeuThrThrValThrGlnSerValThrGlyIleSerArgAsnValSerThrAsn 341  
Db 901 GAAATTTTAAACACAGTACAGATCAGTACAGTATTTCAAGAAATGTAACACTAAT 960  
QY 342 ValPhePheLysGlnHisAspTyrIleIleGluPhePheAspTyrThrThrValIleLys 361  
Db 961 GTGTTCTTCAAGCAACATGATTAATCATCATTTGATTTTGTATTAATCACTGCTTGAAG 1020  
QY 362 ProSerLeuAsnPheThrAlaThrValIleValThrArgAlaAspGlyAsnGlnLeuThr 381  
Db 1021 CCACTTCTCAACTTTCACAGCCACTGTGAGTAACTGCTGCTGATGAGCAACCAACTGACT 1080  
QY 382 LeuGluGluArgArgAsnAsnValIleThrValThrGlnArgAsnTyrThrGluTyr 401  
Db 1081 CTGAAGAAAGAAAGAAATATATGATCATATACATGACAGTACAGAGAAATCTATCGAGTAC 1140  
QY 402 TrpSerGlySerAsnSerGlyAsnGlnLysMetGluAlaValGlnLysIleAsnTyrThr 421  
Db 1141 TGAGCGGAGCTTAACAGTGAATGAAATGAAATGAAAGCTGTTGAAATTAATTAATTA 1200  
QY 422 ValProGlnSerGlyThrPheLysIleGluPheProIleuGluLysSerSerGluLeu 441  
Db 1201 GTCCCCCAAGTGAACCTTTTAAGATTTGAATTCCTCAATCTTGAGAGATTCAGTACGTA 1260  
QY 442 GlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSerLeuPheLys 461  
Db 1261 CAGTTGAAGGCTATTTCTTGGTAGTAAAGTAGACATGGCAGTCACTGCTTTAAG 1320  
QY 462 SerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLysValGlySer 481  
Db 1321 TCTCTTAGTAAAGCATACATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1380  
QY 482 ProPheGluLeuValIleSerGlyAsnLysArgLeuLysGluLeuSerTyrMetValIle 501  
Db 1381 CCTTTGAGTTGGTGTGTTAGTGGCAACAAACGATTAAGAGTTAAGCTTATAGTAGTA 1440  
QY 502 SerArgGlyGlnLeuValAlaValGlyLysGlnAsnSerThrMetPheSerLeuThrPro 521  
Db 1441 TCCAGGGGACAGTTGGGCTGTAGGAAACAAATTTCAACATGTTCTCTTAAACCA 1500  
QY 522 GluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGluAspAspGlyGlu 541  
Db 1501 GAAATTTCTTGAGCTCCAAAGCCGTGTAAATGTGTATTAATTAAGATGATGGGAA 1560  
QY 542 IleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLysLeu 561  
Db 1561 ATTATAGTATGATTTCTAATAATTCCTGTTCACTGTTTATTAATAATAGTAAAGCTA 1620  
QY 562 TyrTrpSerLysValIleLysAlaGluProSerGluLysValSerLeuArgIleSerValThr 581  
Db 1621 TATTTGAGTAAATGAAAGTGAACCATCGAAGAAAGTCTCTTTGAGATCTGTGTACA 1680  
QY 582 GlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsnAla 601  
Db 1681 CAGCTGACTCCATGATGTTGAGTGTGATGTCGTTGAACAAAGTGTGATGATGATGATGCC 1740

QY 602 SerAspIleThrMetGluAsnValValHisGluLeuGluLeuTyrAsnThrGlyTyr 621  
 Db 1741 TCTAATGATATTACAAATGGAATATGTCATGATGGAATTTATTAACAGAGATAT 1800  
 QY 622 TyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTyrValLeu 641  
 Db 1801 TATTTAGCGATGTCATGAAATCTTTTGGAGCTTTCCAGAAATGGAGACTCGGGGATTTG 1860  
 QY 642 ThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyrAspAsnAlaGluTyrAla 661  
 Db 1861 ACGAGTGCACAACTCCAGAGATTTATATGATGGTTTATGACAAATCAAAATAGCT 1920  
 QY 662 GluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeuGly 681  
 Db 1921 GAGAGTTTATGAGAGAAATAGAGACATATTGATGATATTCATGACTTTCTTTGGGT 1980  
 QY 682 SerSerProHisValArgLysHisPheProGluTyrTrpIleTyrLeuAspThrAsnMet 701  
 Db 1981 AGCAGTCCACATGTCGAAAGCATTTTCCAGAGACTTGGATTGGCTAGACCAACATG 2040  
 QY 702 GlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIleThrSerTrpVal 721  
 Db 2041 GGTTCAGGATTTTCCAGAAATTTGAAAGTACTGATCTGATCTATACCTTTCTTTGGGTG 2100  
 QY 722 AlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrProValGlu 741  
 Db 2101 GCTACTGGTTTGTGATCTCTAGAGACCTGGGTCTTGGACTAACAACACTCTCAAGGAG 2160  
 QY 742 LeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArgGly 761  
 Db 2161 CTCCAAGCCTTCCAACTATTTTCAATTTTGTGAATCTTCCCTACTCTGTATCAGAGGT 2220  
 QY 762 GluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrGluValLys 781  
 Db 2221 GAAAGATTTGCTTGGAAATTAATCAATTAATTTGAAGAGTCCACTAGGTAAAG 2280  
 QY 782 ValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsnAla 801  
 Db 2281 GTAATCATTTGAGAAAGTACAAATTTGATATCTTAATGACTTCAATTAATTAATGCTC 2340  
 QY 802 ThrGlyHisGlnGlnThrLeuLeuValProSerGluAspGlyValaThrValLeuPhePro 821  
 Db 2341 ACAGGCCACACAGACACCTTCTGGTCCAGATGAGAGATGGGCAACTGTCTTTTCCC 2400  
 QY 822 IleArgProThrHisLeuGlyGluIleProIleThrValThrAlaLeuSerProThrAla 841  
 Db 2401 ATCAGGCCAACATCTGGGAGAAATCTCTATCACAGTCCAGCTCTTTCACCCACTGCT 2460  
 QY 842 SerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIleGluLysSerTyrSer 861  
 Db 2461 TCTGATGCTGTACCCAGATGATTTTATGTAAGGCTGAAAGAAATGAAATCAATTTCA 2520  
 QY 862 GlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeuSer 881  
 Db 2521 CAATCATCTTATTTAGACTTGAAGTACATAGAGTACAGAGTAACTTGAAGT 2580  
 QY 882 PheSerPheProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIleGly 901  
 Db 2581 TTCTAATTTCTCCTAATACAGTGCAGTGCAGTGAAGATTCAGATCTCAATTTGA 2640  
 QY 902 AspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGlyCys 921  
 Db 2641 GATGTTCTTTGGTCTTCCATCAATGAGCTTACCTATTCGATGCTTATAGGCTGT 2700  
 QY 922 GlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThrLys 941  
 Db 2701 GGTGAACAGAACATGATTAATTTTGTCTCAAAATATTACATTTTGGATTATCTGAATAA 2760  
 QY 942 LysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPheMetArgGlnGlyTyr 961  
 Db 2761 AAGAAACAACCTGACAGATTAATTTGAAGAAAGCTCTTTCAATTAATGAGGCAAGTTAC 2820

QY 962 GluArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyrAsp 981  
 Db 2821 CAGAGAGAACTTCTCTATCAGAGGGAGATGGCTCTTCAAGTCTTTGGGAAATTAATGAC 2880  
 QY 982 ProSerGlySerThrTrpLeuSerAlaPheValLeuArgCysPheLeuGlnAlaAspPro 1001  
 Db 2881 CTTCTGGAGACACTGGGTGTGTGACCTTTGTTTAAAGATGTTTCTTGAAGCCATGCT 2940  
 QY 1002 TyrIleAspIleAspGlnAsnValLeuHisArgThrTyrTrpTrpLeuLysGlyHisGln 1021  
 Db 2941 TACATGATTTGATCAGAAATGTGTACACAGACATACACTTGGCTTAAAGACATCAG 3000  
 QY 1022 LysSerAsnGlyGluPheTrpAspProGlyArgValIleHisSerGluLeuGlnGlyGly 1041  
 Db 3001 AAATCCAACGGTGAATTTTGGATCCAGAGAGATGATTCATTAAGTCAAGCTTCAAGTGGC 3060  
 QY 1042 AsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArgLys 1061  
 Db 3061 AATAAAGTCCAGTAACCTTAACAGCTTATTTGTAACCTTCTCTCTGGGATATGAAAG 3120  
 QY 1062 TyrGlnProAsnIleAspValGlnGluSerIleHisPheLeuGlnSerGluPheSerArg 1081  
 Db 3121 TATCAGCTTAACATTTGATGTGCAGAGCTATCCATTTTGGAGTCTGAATTCAGTGA 3180  
 QY 1082 GlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGlySer 1101  
 Db 3181 GGAATTTCAACAATTAATTAATCTAGCTTATTAATTAATGATGTCATGCTGGAGAGT 3240  
 QY 1102 ProLysAlaLysGluAlaLeuAsnMetLeuThrTrpArgAlaGluGlnGluGlyMet 1121  
 Db 3241 CCTAAGCGGAAGGAGCTTGAATATGCTGACTTGGAGAGCAAGAACAAAGGTGGCAGT 3300  
 QY 1122 GlnPheTrpValSerSerGluSerLysLeuSerAspSerTrpGlnProArgSerLeuAsp 1141  
 Db 3301 CAATCTGGGTGTCAACAGCTCAAGCTTCTGACTCTGGCAGCCAGCTCTCTGGAT 3360  
 QY 1142 IleGluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGluGly 1161  
 Db 3361 ATTGAAGTTGACGCTTAATGACCTGCTCACACTTTCATCAATTTCAAGCTTCAAGGGA 3420  
 QY 1162 IleProIleMetArgTrpLeuSerArgGlnArgAsnSerLeuGlyPheAlaSerThr 1181  
 Db 3421 ATCCCAATTAATGAGGTGCTTAAGCAGCAAGAAATACCTTGGGTGGTTCATCTACT 3480  
 QY 1182 GlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAlaAlaLeuMetAsnThrGlu 1201  
 Db 3481 CAGATTAACCTGTGTGCTTAAAGCTGTCTGAATTTGAGAGCCCTAATGAAATACAGA 3540  
 QY 1202 ArgThrAsnIleGlnValThrValThrGlyProSerSerProSerPro----- 1217  
 Db 3541 AGGACAAATATCCAAAGTACCCGTGACGGGGCTAGCTCACAAAGTCTGTAAAGTTCTG 3600  
 QY 1218 -----LeuAlaValValGlnProMet 1224  
 Db 3601 ATTGACACACAAACCGCTTACTCTTCCAGACAGAGCTTGTGCTGAGCAAGT 3660  
 QY 1225 AlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValTyr 1244  
 Db 3661 GCAGTTAATATTTCGCAAAAGTGTGGATTTGTATTTGTACGCTCAAGTGTGTATAT 3720  
 QY 1245 AsnValLysAlaSerGlySerSerArgArgArgArgSerIleGlnAsnGlnAlaPhe 1264  
 Db 3721 AATGGAAGGCTTCTGGGTCTTCTTGAAGAGAGAGATTTATCCAAATCCAGAGGCTTT 3780  
 QY 1265 AspLeuAspValAlaValLysGluAsnLysAspAsnLeuAsnHisValAspLeuAsnVal 1284  
 Db 3781 GATTAGATGTTGCTGTAAAGAAATTAAGATGATCAATCAATGATGATTTGAATGTG 3840  
 QY 1285 CysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluValAsnLeuLeu 1304  
 Db 3841 TGTACAACTTTTGGGGCCGGGTGAGAGTGGCATGCTCTTAAGAGATTAACCTATTA 3900  
 QY 1305 SerGlyPheMetValProSerGluAlaIleSerLeuSerGluThrValLysValGlu 1324



|||||  
Db AGTGGCTTATAGCGCCCTTCAGAACATTTCTCTGAGCGAGACAGAGAAAGTGGAA 3960  
Qy 1325 TyrAspHisGlyLysIleuAsnLeuTyrIleuAspSerValIangIuThnInPheCysVal 1344  
Db 3961 TATATCATGTGAAACCTCACTTATTAATTCGTGAATGAAACCCAGTTTGTGTT 4020  
Qy 1345 AsnIleProAlaValArgAsnPhelYsValSerAnThnInPheAlaSerValSerIle 1364  
Db 4021 AATATTCCTGCTGTGAGAACTTTAAAGTTCAATATCCAAAGATGCTTCACTCATA 4080  
Qy 1365 ValAspTyrTyrGluProArgArgIleAlaValArgSerTyrAsnSerGluValLysLeu 1384  
Db 4081 GTGATTACTATGAGCCAAAGAGACAGCGCGTGAGAAATTCAACTGTGAAGTGAAGCTG 4140  
Qy 1385 SerSerCysAspLeuCysSerAspValGlnGlyCysArgProCysGluAspGlyVal 1403  
Db 4141 TCTTCCTGTGACCTTGTGAGTGAATGCCAGGCGTCCCTCTGTGAGATGAGCT 4197  
RESULT 14  
ABQ79968 standard; cDNA: 3535 BP.  
ID ABQ79968  
AC ABQ79968;  
DT 23-DEC-2002 (first entry)  
XX  
XX Human CD109 K15 protein encoding cDNA.  
DE  
XX CD109; alpha2 macroglobulin; thioester; cerebroprotective; cardiant;  
KW immunosuppressive; haemostatic; anticoagulant; thrombolytic; human;  
KW cardiovascular; vasotropic; gene therapy; CD109 K15; gene; ss.  
XX  
OS Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FH 113..4450  
FT /\*tag= a  
FT /product= "CD109 K15"  
XX  
XX MO200270696-A2.  
XX  
XX 12-SEP-2002.  
XX  
XX 07-MAR-2002; 2002WO-CA00292.  
XX  
XX 07-MAR-2001; 2001US-273814P.  
XX  
XX (SCHU/) SCHUH A.  
XX (SUTH/) SUTHERLAND R D.  
XX  
XX Schuh A, Sutherland RD;  
XX  
XX WPI; 2002-713450/77.  
XX P-PSDB; ABB82169.  
XX  
XX  
XX New CD109 nucleic acids and polypeptides, useful in gene therapy,  
XX particularly for treating strokes, myocardial infarctions, thrombosis,  
XX thrombocytopenia, autoimmune diseases, or organ or bone marrow  
XX transplantation -  
XX  
XX Claim 1; Fig 4a; 156pp; English.  
XX  
XX The invention relates to isolated nucleic acid molecules encoding CD109  
XX polypeptides. These nucleic acid molecules include the human cDNA  
XX sequences comprising CD109 K1, CD109 K1-H7, CD109 K15 or their variants.  
XX CD109 is a novel member of the alpha2 macroglobulin (alpha2M)/C3, C4, C5  
XX family of thioester-containing proteins. The CD109 polypeptides can be  
XX expressed by standard recombinant methodology. The CD109 nucleic acid,  
XX CD109 polypeptide or its mimetic, or the anti-CD109 antibody is useful  
XX for treating or detecting a disease or disorder, e.g. conditions  
XX associated with endothelial activation, platelet activation, activation  
XX of the coagulation or fibrinolytic systems, activation of T lymphocytes

CC and of the complement system, quantitative or qualitative abnormalities  
CC of platelet function, increased or impaired platelet aggregation and  
CC activation, increased or impaired activation of the coagulation and/or  
CC fibrinolytic systems, or impaired or increased immune activation. These  
CC are also useful for treating cardiovascular disorders, stroke, myocardial  
CC infarction, thrombosis, embolism, peripheral vascular disease,  
CC thrombocytopenia, thrombocythemia, autoimmune diseases, organ  
CC transplantation, or bone marrow transplantation. The CD109 nucleic acids  
CC or their antisense nucleotide sequence are useful in gene therapy for  
CC treating these conditions. The present sequence represents the human  
CC CD109 K15 cDNA sequence.  
XX  
SQ Sequence 3535 BP; 1108 A; 675 C; 731 G; 1021 T; 0 other;  
Alignment Scores:  
Pred. No.: 0 Length: 3535  
Score: 5470.00 Matches: 1062  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.91% Mismatches: 0  
Query Match: 74.44% Indels: 0  
DB: 24 Gaps: 0  
US-10-020-095-4 (1-1428) x ABQ79968 (1-3535)  
Qy 1 MetGlnGlyProProLeuLeuThnAlaAlaIleuLeuCysValCysThnAlaAlaLeu 20  
Db 113 ATGAGAGGCCACCGCTCTGACCGCGCCACCTCTGCGTGTGACCGCGCGCTG 172  
Qy 21 AlaValAlaProGlyProArgPheLeuValThnAlaProGlyLysIleLeaProGlyGly 40  
Db 173 GCGGTGCTCCCGGCGCTCGGTTTGTGTGACAGCCCGAGAGATCATCAGCGCCGAGGA 232  
Qy 41 AsnValThrIleGlyValGluLeuLeuGluHisCysProSerGlnValThnValLysAla 60  
Db 233 AATGTGACTATTTGGGGTGGAGCTTCTGGACATCTGCCCTTCAAGGTGATCTGAAGCG 292  
Qy 61 GluLeuLeuLysThnAlaSerAsnLeuThnValSerValLeuGluValGluGlyValPhe 80  
Db 293 GAGCTGCTCAAGACAGATCAAACTCATCTGTCTGTCTGAGAGCAGAAAGTCTTT 352  
Qy 81 GluLysGlySerPheLysThrLeuThnLeuProSerLeuProLeuAsnSerAlaAspGlu 100  
Db 353 GAAAGAGCTCTTTTAAAGACACTTACTCTTCATCACTTACCTTGAGACAGTGCAGAGAG 412  
Qy 101 IleTyrGluLeuArgValThnGlyArgThnGluAspGluIleLeuPheSerAsnSerThr 120  
Db 413 ATTATGAGCTACGTGTACCGAGCGTACCCAGATGAGATTTATTTCTTAATAGTACC 472  
Qy 121 ArgLeuSerPheGluThnLysArgIleSerValPheIleGlnThnAspLysAlaLeuTyr 140  
Db 473 CGCTTATCATTTGAGACCAAGAGATATCTGTCTTCAATCAACAGCAAGCCTTATAC 532  
Qy 141 LysProLysGlnGluValLysPheArgIleValThnLeuPheSerAspPheLysProTyr 160  
Db 533 AAGCCAAAGCAAGAGAGGAAATTCGATTTGATCACTCTTCAGATTTTAAAGCCCTTAC 592  
Qy 161 LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnIntProLeu 180  
Db 593 AAAACCTCTTTAAACATTCATTTAAGAGCCCAAAATTTGATCCAAAGTGTG 652  
Qy 181 SerGlnIleSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIle 200  
Db 653 TCACAAAGATGATCTGTGAGATCATTTCCAAACTTTTACGATCTTCCATCCATA 712  
Qy 201 LeuGlyAspTyrSerIleGlnValGlnValAsnAspGlnThnTyrTyrGlnSerPheGln 220  
Db 713 CTGTGTACGTGCTATTCATCAAGTTCAAGTGAAGACAGACATATATCATCATTTTCAG 772  
Qy 221 ValSerGluTyrValLeuProLysPheGluValThnLeuInThnProLeuTyrCysSer 240  
Db 773 GTTTCAAATATGATATTAACAAATTTGAAGTGACTTGTGAGACCATATATATGTTCT 832  
Qy 241 MetAsnSerLysHisLeuAsnGlyThrIleThnAlaLysTyrThrTyrGlyLysProVal 260

Db 833 ATGAATTCCTAAGCATTTAAATGATGACATCAAGCAAGATATACATATGGGAAGCCAGAGT 892  
Qy 261 LysGlyAspValThrLeuThrPheLeuProLeuSerPheTrpGlyIleuValAsnIle 280  
Db 893 AAGAGAGAGTAAAGCTTACATTTTACCTTATCTCTTTGGGAAAGAAAGAAATATT 952  
Qy 281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluMetLys 300  
Db 953 ACAAAAACATTTAAGATTAATGATCTGCAAACTCTCTTTATGATGAAGATGAATA 1012  
Qy 301 AsnValMetAspSerSerSerSerSerGlyLeuSerGluIleuAspLeuSerSerProGlyPro 320  
Db 1013 AATGATATGATTCCTTCAATATGACATTCCTGATACCTGATCTATCTCCCTGACACA 1072  
Qy 321 ValGluIleLeuThrThrValThrGluSerValThrGlyIleSerArgAsnValSerThr 340  
Db 1073 GTRGAATTTTAAACACAGTACAGAAATCACTTACAGATTTTCAAGAAATGTAAAGCACT 1132  
Qy 341 AsnValPhePheLysGlnHisAspTyrIleIleGluPhePheAspTyrThrThrValLeu 360  
Db 1133 AATGTCTCTTCAAGCAACATGATTAATCATTTGAGTTTGTATTAATCTACTGTCTTG 1192  
Qy 361 LysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsnGluLeu 380  
Db 1193 AAGCCATCTCTCAACTCAACAGCACAGTGAAGATACTGCTGATGCAACCAACTG 1252  
Qy 381 ThrLeuGluGluArgArgAsnValValIleThrValThrGlnArgAsnTyrThrGlu 400  
Db 1253 ACTCTTGAAAGAAAGAAATATATGATGATACATGACAGTACAGAAATCTATCTGAG 1312  
Qy 401 TyrTrpSerGlySerAsnSerGlyAsnGlnLysMetGluAlaValGlnLysIleAsnTyr 420  
Db 1313 TACTGAGCGGAGCTTAACAGTGAATTCAGAAATGGAAGCTGTCAGAAATTAATTAT 1372  
Qy 421 ThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGluAspSerSerGlu 440  
Db 1373 ACTGTCGCCCAAGTGAACCTTTAAGATGAATTCCAATCTCGAGAGATTCAGTGTAG 1432  
Qy 441 LeuGlnLeuLysAlaThrPheLeuGlySerLysSerSerMetAlaValHisSerLeuPhe 460  
Db 1433 CTACAGTTAAGGCTTAATTCCTTGGTAGAAAGATGACATGCACTCATCTGTTT 1492  
Qy 461 LysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLysValGly 480  
Db 1493 AAGTCTCCATAGTAAAGCATACATCCAACTAAACAAAGATGAATAATTAAGTGGGA 1552  
Qy 481 SerProPheGluLeuValValSerGlyAsnLysArgLeuLysGluLeuSerTyrMetVal 500  
Db 1553 TCGCCTTTTGAAGTGGTGGTGGTGGCAACAAACGATTGAGAGTTAAGCTATATGTA 1612  
Qy 501 ValSerArgGlyGlnLeuValAlaValGlyLysGlnAsnSerThrMetPheSerLeuThr 520  
Db 1613 GATTCAGAGGAGACAGTTGGTGGTGAAGAAACAAATTCACAAATGTTCTCTTAAACA 1672  
Qy 521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGluAspAspGly 540  
Db 1673 CCAAGAAATTTCTTGGACTCCAAAGCTGTGTATGTTGTATTAATTAAGATGAAGG 1732  
Qy 541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560  
Db 1733 GAATATATAGTATGTTCTTAAATAATTCCTGTTCACTGTTTAAATAATAGATTAAG 1792  
Qy 561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580  
Db 1793 CTATATATGAGTAAAGTGAAGCTGAACCATCTGAAGAAAGTCTCTTGGAGTCTGTGG 1852  
Qy 581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600  
Db 1853 AACAGCCTGACCTCCATAGTGGGATTTGATGCTGTTGACAAAGTGTGAATCTGAAGAT 1912  
Qy 601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeuGluLeuTyrAsnThrGly 620

Db 1913 GCCTTAATGATATTAATCAATGGAATAATGTGCTCATGATGTTGAACTTTATTAACACAGA 1972  
Qy 621 TyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTrpVal 640  
Db 1973 TATTAATTAAGCATGTTCTATGAATTTCTTTGCAAGTCTTTCAGAAATGTGAATCTGGGTA 2032  
Qy 641 LeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyrAspAsnAlaGluTyr 660  
Db 2033 TTGACAGATGACAACTCCAGAAAGATTTATATGATGGTGTATATGACATGACAGATAT 2092  
Qy 661 AlaGluArgPheMetGluGluLysAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680  
Db 2093 GCTGAAGGTTTATGAGAGAAATGAAGACATATGTATGATATTCATGACTTCTTCTTGG 2152  
Qy 681 GlySerSerProHisValArgLysHisPheProGluThrTrpIleTrpLeuAspThrAsn 700  
Db 2153 GGTAGCAGTCCACATGTCGAAAGCATTTTCCAGAGACTTGAAATTTGGCTAAGACACCAAC 2212  
Qy 701 MetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIleThrSerTrp 720  
Db 2213 ATGGGTTACAGATTTTACCAAGAAATTTGAAGTAACTGTAACCTGATCTATCACTCTTGG 2272  
Qy 721 ValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrProVal 740  
Db 2273 GTGGCTACTGTTTGTGTGATCTCTGAGGACCTGGGCTTGGACTTAACAACACTCCAGTG 2332  
Qy 741 GluLeuGlnAlaPheGluProPhePheIlePheLeuAsnLeuProTyrSerValIleArg 760  
Db 2333 GAGCTCCAGCCTTCCAACTTTTTCATTTTGTGAATCTTCCACTCTGTTATCAACA 2392  
Qy 761 GlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrGluVal 780  
Db 2393 GGTGAAGATTTGCTTGGAAATTAATCAATTAATTTGAAGATCCCACTGAGGTT 2452  
Qy 781 LysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSerGluLeuAsn 800  
Db 2453 AAGGTATCATTTGAGAAAGATGAATAATTTGATTTATCTTAATGACTTCAATGAATAAT 2512  
Qy 801 AlaThrGlyHisGlnGluThrLeuLeuValProSerGluAspGlyValaThrValLeuPhe 820  
Db 2513 GGCACAGGCCACACACAGACCTTCTGTTCCAGTGAAGAGTGGGCAACTGTCTTTT 2572  
Qy 821 ProIleArgProThrHisLeuGlyGluIleProIleThrValThrAlaLeuSerProThr 840  
Db 2573 CCCATCAGGCCAACATCTGGGAGAAATCTCTATCACAGTACAGCTCTTCAACCCACT 2632  
Qy 841 AlaSerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIleGluLysSerTyr 860  
Db 2633 GCTTCTGATGCTGTCAACCAAGATTTTATGTAAGGCTGAAGAGATGAATAATCATAT 2692  
Qy 861 SerGlnSerIleLeuLeuAspLeuThrAspAsnAlaGluLeuGlnSerThrLeuLysThrLeu 880  
Db 2693 TCACATCCATCTTATTAAGACTGTGACTGACATAGGCTTACAGATACCTGAAACCTTGG 2752  
Qy 881 SerPheSerPheProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIle 900  
Db 2753 AGTTTCTCTTTCTCTCTATACAGTGACTGGCAAGTGAAGAGCTTCAGATCCAGATAT 2812  
Qy 901 GlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGly 920  
Db 2813 GGAGATGTTCTTGGTCTTCATCAATGAGCTTACCTCATGATTCGATGCTTATGAGC 2872  
Qy 921 CysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThr 940  
Db 2873 TGTGGTGAACAAACATGATTAATTTGTCTCAAAATATTTAATTTGATTAATCTGACT 2932  
Qy 941 LysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPheMetArgGlnGly 960  
Db 2933 AAAAGAAACACATGACAGTATATTTGAAGAAAGAAAGCTCTTCAATTAATGAGCAAGT 2992  
Qy 961 TyrGlnArgGluLeuLeuTyrGlnArgLysAspGlySerPheSerAlaPheGlyAsnTyr 980  
Db 2993 TACCAAGAGAACTCTCATACAGAGGAGATGAGTCTTTCAGAGCTTTTGGGAATATAT 3052

QY 981 AspProSerGlySerThrTyrLeuSerAlaPheValLeuArgCysPheLeuGluAlaAsp 1000  
Db 3053 GACCTTCTGGAGACATGGTGTGCAGCTTTGTTTAAATGTTCTTGAAAGCCGAT 3112  
QY 1001 ProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTyrLeuGlyHis 1020  
Db 3113 CCTTACATAGATATGATGACAAATGTGTTCACAGAACATACACTTGCTTAAAGGACAT 3172  
QY 1021 GlnTyrSerAsnGlyGluPheThrAspProGlyArgValIleHisSerGluLeuGlnGly 1040  
Db 3173 CAGAAATCAACGGTGAATTTTGGATCCAGAGAGATTCATAGTACGCTTCAGAGCT 3232  
QY 1041 GlnAsnTyrSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArg 1060  
Db 3233 GGCATATAAGTCCAGTAACTTACAGCTTATATGTAATTTCTCTCTCGGATATAGA 3292  
QY 1061 LysTyrGln 1063  
Db 3293 AAGTATCAG 3301  
RESULT 15  
ABQ79969  
ID ABQ79969 standard; cDNA, 3535 BP.  
AC ABQ79969;  
XX  
XX  
XX 23-DEC-2002 (first entry)  
DT  
XX  
XX Human CD109 K15 variant protein encoding cDNA.  
DE  
XX  
XX CD109; alpha2 macroglobulin; thioester; cerebroprotective; cardiant;  
KW immunosuppressive; haemostatic; anticoagulant; thrombolytic; human;  
KW cardiovascular; vasotropic; gene therapy; CD109 K15; variant; gene; ss.  
XX  
XX Homo sapiens.  
OS  
XX  
XX Key Location/Qualifiers  
FH 113..4450  
FT CDS /tag= a  
FT /product= "CD109 K15 variant"  
FT  
XX  
XX WO200270696-A2.  
XX  
XX 12-SEP-2002.  
XX  
XX 07-MAR-2002; 2002WO-CA00292.  
XX  
XX 07-MAR-2001; 2001US-273814P.  
XX  
XX (SCHU// SCHUH A.  
XX (SUTR// SUTHERLAND R D.  
XX  
XX Schuh A, Sutherland RD;  
XX  
XX WPI; 2002-713450/77.  
XX  
XX P-PSDB; ABB82170.  
XX  
XX  
XX New CD109 nucleic acids and polypeptides, useful in gene therapy,  
XX particularly for treating strokes, myocardial infarctions, thrombosis,  
XX thrombocytopenia, autoimmune diseases, or organ or bone marrow  
XX transplantation -  
XX  
XX  
XX Claim 1; Fig 4b; 156pp; English.  
XX  
XX The invention relates to isolated nucleic acid molecules encoding CD109  
XX polypeptides. These nucleic acid molecules include the human cDNA  
XX sequences comprising CD109 K1, CD109 K1-H7, CD109 K15 or their variants.  
XX CD109 is a novel member of the alpha2 macroglobulin (alpha2m)/C3, C4, C5  
XX family of thioester-containing proteins. The CD109 polypeptides can be  
XX expressed by standard recombinant methodology. The CD109 nucleic acid,  
XX CD109 polypeptide or its mimetic, or the anti-CD109 antibody is useful  
XX for treating or detecting a disease or disorder, e.g. conditions

CC associated with endothelial activation, platelet activation, activation  
CC of the coagulation or fibrinolytic systems, activation of T lymphocytes  
CC and of the complement system, quantitative or qualitative abnormalities  
CC of platelet function, increased or impaired platelet aggregation and/or  
CC activation, increased or impaired activation of the coagulation and/or  
CC fibrinolytic systems, or impaired or increased immune activation. These  
CC are also useful for treating cardiovascular disorders, stroke, myocardial  
CC infarction, thrombosis, embolism, peripheral vascular disease, organ  
CC thrombocytopenia, thrombocytopenia, autoimmune diseases, organ  
CC transplantation, or bone marrow transplantation. The CD109 nucleic acids  
CC or their antisense nucleotide sequence are useful in gene therapy for  
CC treating these conditions. The present sequence represents the human  
CC CD109 K15 variant cDNA sequence.  
XX  
SQ Sequence 3535 BP; 1107 A; 676 C; 731 G; 1021 T; 0 other;  
Alignment Scores:  
Pred. No.: 0 Length: 3535  
Score: 5461.00 Matches: 1061  
Percent Similarity: 99.91% Conservative: 1  
Best Local Similarity: 99.81% Mismatches: 1  
Query Match: 74.32% Indels: 0  
DB: 24 Gaps: 0  
US-10-020-095-4 (1-1428) x ABQ79969 (1-3535)  
QY 1 MetGlnGlyProProLeuLeuThrAlaAlaHisLeuLeuCysValCysThrAlaAlaLeu 20  
Db 113 ATGACGGGCCCCACCGCTCTGACCGCCGACCTCTCTGCTGTGCACCGCGCGCTG 172  
QY 21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGlyGly 40  
Db 173 GCGGTGGCTCCGCGGCTCGGTTTGTGTGACAGCCCGAGGATCATACAGCCCGAGGA 232  
QY 41 AsnValThrIleGlyValGluLeuLeuGluHisCysProSerGlnValThrValValAla 60  
Db 233 AATGTGACTATTTGGGGGAGCTTCGGAACATCGCTTCACAGGTGATGGAAGCG 292  
QY 61 GluLeuLeuThrAlaSerAsnLeuThrValSerValLeuGluAlaGluValPhe 80  
Db 293 GAGCTGCTCAGACAGCATTAACCTCCTGCTGTCTGTGAGAGCAGAGAGCTTT 352  
QY 81 GluLeuGlySerPheLeuThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGlu 100  
Db 353 GAAAGGCTCTTTTAAAGCACTTACTCTTCATCACTACTACTCTGAAACAGTGAATGAG 412  
QY 101 IleTyrGluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThr 120  
Db 413 ATTTATGAGCTACGTGTAAACCGGACGTAACAGATGAGATTTTATTTCTTAATAGTACC 472  
QY 121 ArgLeuSerPheGlnThrLysArgIleSerValPheIleGlnThrAspLysAlaLeuTyr 140  
Db 473 CGCTTATCATTTAGACCAAGAAATATCTGCTTCAATTCACAGCAAGGCTTATAC 532  
QY 141 LysProLeuGlnIleValLysPheArgIleValThrLeuPheSerAspPheLeuProTyr 160  
Db 533 AAGCCAAAGCAAGAGAGAGTTTCCGATTTGTAACCTCTTCAGATTTTAAAGCTTAC 592  
QY 161 LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnIleThrLeu 180  
Db 593 AAAACCTCTTAAACATTCATTAAGAGACCCCAAAATTCAAATTTGATCCAAACGTGTTG 652  
QY 181 SerGlnGlySerAspLeuGlyValIleSerThrPheGlnLeuSerSerHisProIle 200  
Db 653 TCACAAACAAAGTATCTGGAGTCATTTCCAAACATTTTACGCTATCTCCATCCATA 712  
QY 201 LeuGlyAspTyrPheIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGln 220  
Db 713 CTGGTGACCTGGCTATTCAGTTCAGTGAAGTGAAGCAAGCAATATATATCATTTTCA 772  
QY 221 ValSerGlnTyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSer 240  
Db 773 GTTTCAGAAATATATATTACCAAAATTTGAAGTACCTTGGACAGACACCATTAATTTGTT 832

QY 241 MetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysThrTyrGlyLysProVal 260  
 DB 833 ATGAATTTCTAAGCATTTAAATGATACCATCAGCGAAAGATATACATATGGAAGCCAGTG 892  
 QY 261 LysGlyAspValThrLeuThrPheLeuProLeuSerPheTyrGlyLysLysAsnIle 280  
 DB 893 AAAGGAGAGCTAACCGCTTACATTTTACCTTTATCCTTTTGGGAAAGAAATATAT 952  
 QY 281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluGluMetLys 300  
 DB 953 ACMAAAACATTTAAGATTAATGATCTGCAAACTTCTTTATATGATAGAAGAGAA 1012  
 QY 301 AsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSerProGlyPro 320  
 DB 1013 AATGTAATGGAATCTTCAATAGCACTTCTGAATACCTGAGTCTATCTTCCCTGGACCA 1072  
 QY 321 ValGluIleLeuThrThrValThrGluSerValThrGlyIleSerArgAsnValSerThr 340  
 DB 1073 GTAGAAATTTTAAACCAAGTGCAGATCAGATCAGTATTCAGGATTTCAAGAAATGTAAAGCACT 1132  
 QY 341 AsnValPhePheLysGlnHisAspTyrIleIleGluPheAspTyrThrThrValLeu 360  
 DB 1133 AATGTTCTTCAAGCAAGTATTAATCATCATGAGTTTGTGATTAATCTACTGCTTG 1192  
 QY 361 LysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsnGluLeu 380  
 DB 1193 AAGCATCTCTCAACTTCAAGCACTGAGCACTGAGAGTAACTCGTCTGATGCAACCAACTG 1252  
 QY 381 ThrLeuGluGluArgArgAsnValValIleThrValThrGlnArgAsnTyrThrGlu 400  
 DB 1253 ACTCTTGAAGAAAGAAATATATGATCATTAACAGTACAGACAGAAACTATACTGAG 1312  
 QY 401 TyrTrpSerGlySerAsnSerGlyAsnGlyLysMetGluAlaValGluMetLysLeuThr 420  
 DB 1313 TACTGAGGCGGANTCAACAGTGAATCAAGAAATCAAGAAAGCTGTCAGAAATTAATTTAT 1372  
 QY 421 ThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGluAspSerSerGlu 440  
 DB 1373 ACTGTCCTCCCAAGTGGAACTTTTAAATGATGTAATCTCCAACTCTGAGAGATTCCTAGTGG 1432  
 QY 441 LeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSerLeuPhe 460  
 DB 1433 CTACAGTTGAAGCGCTTAATTCCTTGTGATGTAAGAAAGTACATGACATGCTGTTT 1492  
 QY 461 LysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLysValGly 480  
 DB 1493 AAGTCTCCATGATGAACATACATCCAACTAAACAAAGATGAAATATTAAGGTGGGA 1552  
 QY 481 SerProPheGluLeuValValSerGlyAsnLysArgLeuLeuLeuSerTyrMetVal 500  
 DB 1553 TCCTCTTTGAGTGGTGGTGTAGTGCACAAACATTAAGAGTTAAAGCTATATGATA 1612  
 QY 501 ValSerArgGlyGlnLeuValAlaValGlyLysGlnAsnSerThrMetPheSerLeuThr 520  
 DB 1613 GATATCAGGGGACAGTTGGTGGCTGTAGGAAACAAATTTCAACATGTTCTCTTTAACA 1672  
 QY 521 ProGluAsnSerThrProLysAlaCysValIleValTyrTyrIleGluAspAspGly 540  
 DB 1673 CCGAAAAATTTGGAATCCAAAGCTGTGTAATGTGATTAATTTAAGATGATGGG 1732  
 QY 541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560  
 DB 1733 GAATTTATTAAGTATGTTCTTAAATAATCTCTGTTCACTGTTTAAAATAATAGATTAAG 1792  
 QY 561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580  
 DB 1793 CTATATTGAGTAAAGTGAAGTGAACATCTGAAGAAAGTCTCTTGAGATCTCTGG 1852  
 QY 581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600  
 DB 1853 AACACGCTGACCTCATAGTGGGATTTAGCTGTGACAAAGATGTGATCTGATGAAT 1912

QY 601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeuGluLeuTyrAsnThrGly 620  
 DB 1913 GCCTTATATGATATTAACAATGGAATAATGTGGTCCATGAGTTGGAACCTTATTAACAAGGA 1972  
 QY 621 TyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuThrVal 640  
 DB 1973 TATTATTTAGGACATGTTTATTAATATTTCTTTTCCAGCTTTCAGGATGTGGACTCGGGCTA 2032  
 QY 641 LeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyrAspAsnAlaGluTyr 660  
 DB 2033 TTGACAGATGCAAACTTACAGAGGATTAATGATGAGTGTGTTATGACAAATGCAGAAAT 2092  
 QY 661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680  
 DB 2093 GCTGAGAGTATTATGAGAAAGAAATGAAGACATATTTGATGATATTCATGACTTTTCTTTG 2152  
 QY 681 GlySerSerProHisValArgLysHisPheProGluThrTrpIleTrpLeuAspThrAsn 700  
 DB 2153 GGTACAGTCCACATGTCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACCCAAAC 2212  
 QY 701 MetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIleThrSerTrp 720  
 DB 2213 ATGGTTCCAGGATTTACCAAGAAATTTGAAGTAACTGATCTGATTTATCACTTCTTGG 2272  
 QY 721 ValAlaThrGlyPheValIleSerGluAspLeuGlyLeuThrThrThrProVal 740  
 DB 2273 GTGGCTAATCGGTTTGTATCTGTAGAGCACTGGGTTCTGAGCTAACACTCACTCACTG 2332  
 QY 741 GluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArg 760  
 DB 2333 GAGCTCCAAAGCCTTCCAAACCAATTTTTCATTTTGTGAATCTTCCCTACTCTGTTATCA 2392  
 QY 761 GlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrGluVal 780  
 DB 2393 GGTGAAGATTTGCTTGTGAATTAATATTAATTAATTAATTAATTAATTAATTAATTAAT 2452  
 QY 781 LysValIleIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsn 800  
 DB 2453 AAGGTATATCTTGAAGAAAGTGAACAAATTTGATATTTCAATGACTTCAAAAGAAATTAAT 2512  
 QY 801 AlaThrGlyHisGlnGlnThrLeuLeuValProSerGluAspGlyValaThrValLeuPhe 820  
 DB 2513 GCCACAGGCCACACAGCAACCTTCTGTGTTCCAGTGAGATGGGGCAACTGTTCTTTT 2572  
 QY 821 ProIleArgProThrHisLeuGlyGluIleProIleThrValThrAlaLeuSerProThr 840  
 DB 2573 CCCATCAGGCCCAACATCTGGAGAAATTCCTATCAACAGTCAAGCTCTTTCACCCACT 2632  
 QY 841 AlaSerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIleGluLysSerTyr 860  
 DB 2633 GCTTGTGATGCTGTCAACCAATGATTTTATTAAGCTGAAGAAATGAAGAAATCAAT 2692  
 QY 861 SerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeu 880  
 DB 2693 TCACATATCATTTATTAATGACTTGACTGACATAGAGTCAAGAGTAAACCTTGAACCTTTG 2752  
 QY 881 SerPheSerPheProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIle 900  
 DB 2753 AGTTTCTAATTTCTCTTAATACAGTGCAGTGGAGTGAAGAGTTCAATCAATCACTGCAAT 2812  
 QY 901 GlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGly 920  
 DB 2813 GGAAGATGTTCTTGCTTCCATCAATAGCTTTAGCTTCATGATTCGGAATGCCCTTAATG 2872  
 QY 921 CysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThr 940  
 DB 2873 TGTGTGAACAGAACATGATTAATTTCTCCAAATATTTTCAATTTTGATTAATCTGACT 2932  
 QY 941 LysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPheMetArgGlnGly 960  
 DB 2933 AAAAAGAAACCACTGACAGATTAATTTGAAGAAAGCTCTTTCATTTATAGAGCAAGGT 2992  
 QY 961 TyrGlnArgGluLeuLeuTyrGlnArgGluAspLysSerPheSerAlaPheGlyAsnTyr 980

```
Db      2993  TACGAGAGAACTTCTCTATCAGAGGAGATGGCTCTTCAGTGTCTTTGGGAATTAT 3052
Qy      981  AspProSerGlySerThrTripleuSerAlaPheValLeuArgCysPheLeuGluAlaAsp 1000
Db      3053  GACCTTCTGGGAGACACTTGGTGTGCACCTTTGTTTAAGATGTTCTTGAAGCCGAT 3112
Qy      1001  ProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTripleuLysGlyHis 1020
Db      3113  CTTACATAGATATTGATCAGAAATGTGTACACAGAACATACACTTGGCTTAAAGGACAT 3172
Qy      1021  GlnLysSerAspGlyGluPheTrpAspProGlyArgValIleHisSerGluLeuGlnGly 1040
Db      3173  CAGAAATCCAACGGTGAATTGGGATCCAGGAGAGTGAATTCATAGTGAAGCTTCAAGGT 3232
Qy      1041  GlyAsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArg 1060
Db      3233  GGCATTAATAAGTCCAGTAACCTTACAGCCTTAATTGTAACCTTCTCTCTGGGATATGA 3292
Qy      1061  LysTyrGln 1063
Db      3293  AAGTATCAG 3301
```

Search completed: January 17, 2004, 01:15:04  
Job time : 983 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: January 16, 2004, 21:02:03 ; Search time 198 Seconds

(without alignments)  
3183.309 Million cell updates/sec

Title: US-10-020-095-4

Perfect score: 7348  
Sequence: 1 MGGPPLTAHLVCVCTAL.....HSSVIFPCFKLYEMELWL 1428

Scoring table: BLOSUM62

Xgapop 10.0	Xgapext 0.5
Ygapop 10.0	Ygapext 0.5
Fgapop 6.0	Fgapext 7.0
Delop 6.0	Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+.p2n.model -DBV=xlp  
-Q=/cg2\_1/uspro\_pool\_p/us10020095/runat.15012004.163823.18719/app.query.fasta.1.1607  
-DB=Issued\_Patents\_NA -QFMT=fastap -SUFFIX=rn1 -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bloms62 -TRANS=human40.cdi  
-LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10020095 @CGN 1.1 118 @runat.15012004.163823.18719 -NCPU=6 -ICPU=3  
-NO MAP -LARGEOUTERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOC  
-DBV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA: \*  
1: /cg2\_6/prodata/2/ina/5A.COMB.seq: \*  
2: /cg2\_6/prodata/2/ina/5B.COMB.seq: \*  
3: /cg2\_6/prodata/2/ina/6A.COMB.seq: \*  
4: /cg2\_6/prodata/2/ina/6B.COMB.seq: \*  
5: /cg2\_6/prodata/2/ina/ECTUS.COMB.seq: \*  
6: /cg2\_6/prodata/2/ina/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1441	19.6	4577	4	US-09-241-606-1
2	1346.5	18.3	4079	4	US-09-016-434-1114
3	973.5	13.2	5211	1	US-08-447-411-1
4	931	12.7	5056	2	US-08-793-126-2
5	931	12.7	5056	3	US-09-132-271-2
6	931	12.7	5067	3	US-09-142-334-23
7	894	12.2	5948	2	US-08-662-227-1
8	894	12.2	5948	4	US-09-017-947-1
9	885	12.0	5924	1	US-08-447-411-44
10	768.5	10.5	4138	1	US-08-447-411-75
11	768.5	10.5	4138	2	US-08-662-227-33
12	768.5	10.5	4138	4	US-09-017-947-33

C	13	245	3.3	18431	4	US-09-221-017B-1090	Sequence 1090, Ap
	14	234	3.2	339	4	US-09-311-352B-1	Sequence 1, Appli
	15	215.5	2.9	750	4	US-09-241-606-3	Sequence 3, Appli
C	16	203.5	2.8	14066	4	US-09-601-198-56	Sequence 56, Appli
	17	185	2.5	549	4	US-09-311-352B-3	Sequence 3, Appli
	18	180	2.4	4026	4	US-09-252-991A-10214	Sequence 10214, A
	19	174.5	2.4	7101	1	US-08-480-604A-9	Sequence 9, Appli
	20	174.5	2.4	7101	2	US-08-405-496A-9	Sequence 9, Appli
	21	174.5	2.4	7101	3	US-08-915-136-9	Sequence 9, Appli
	22	174.5	2.4	7101	4	US-08-957-310-9	Sequence 9, Appli
	23	174.5	2.4	7101	4	US-10-011-366-9	Sequence 9, Appli
	24	171	2.3	9873	4	US-09-328-352-1360	Sequence 1360, Ap
	25	165	2.2	28473	4	US-08-961-527-83	Sequence 83, Appli
	26	164.5	2.2	4977	4	US-09-071-035-257	Sequence 257, App
	27	164.5	2.2	4977	4	US-09-071-035-261	Sequence 265, App
	28	164.5	2.2	4977	4	US-09-071-035-265	Sequence 265, App
	29	164.5	2.2	13993	4	US-09-220-132-20	Sequence 20, Appli
	30	163.5	2.2	14070	4	US-09-108-006C-2	Sequence 2, Appli
	31	163	2.2	37948	3	US-09-251-645-11	Sequence 11, Appli
C	32	161	2.2	3312	4	US-09-601-198-58	Sequence 58, Appli
	33	161	2.2	7286	4	US-08-793-273C-1	Sequence 1, Appli
	34	161	2.2	7286	5	PCT-US95-11684-1	Sequence 1, Appli
	35	160	2.2	5629	4	US-09-453-702B-243	Sequence 243, App
C	36	157.5	2.1	15016	4	US-09-601-198-60	Sequence 60, Appli
C	37	156	2.1	3636	4	US-08-961-527-78	Sequence 78, Appli
	38	156	2.1	30549	4	US-09-134-001C-322	Sequence 322, App
	39	155.5	2.1	9510	4	US-09-453-702B-256	Sequence 256, App
	40	154	2.1	9171	1	US-08-038-682-5	Sequence 5, Appli
	41	154	2.1	9171	1	US-08-302-832-5	Sequence 5, Appli
	42	154	2.1	9171	1	US-08-530-198-5	Sequence 5, Appli
	43	154	2.1	9171	2	US-08-469-880-5	Sequence 5, Appli
	44	154	2.1	9171	2	US-08-728-470-5	Sequence 5, Appli
	45	154	2.1	9171	2	US-08-617-697-5	Sequence 5, Appli

#### ALIGNMENTS

RESULT 1  
US-09-241-606-1  
; Sequence 1, Application US/09241606  
; Patent No. 6472140  
; GENERAL INFORMATION:  
; APPLICANT: Tanzi, Rudolph E.  
; APPLICANT: Kovacs, Dora  
; APPLICANT: Saunders, Aleister J.  
; TITLE OF INVENTION: Alpha-2-Macroglobulin Therapies and Drug Screening Methods for  
; FILE REFERENCE: 0609.4460003  
; CURRENT APPLICATION NUMBER: US/09/241,606  
; CURRENT FILING DATE: 1999-02-02  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 4577  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: sig\_peptide  
; LOCATION: (44)..(112)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (44)..(4468)  
; FEATURE:  
; NAME/KEY: mat\_peptide  
; LOCATION: (113)..(4468)  
; US-09-241-606-1

Alignment Scores:  
Pred. No.: 1,496-142  
Score: 1441.00  
Percent Similarity: 46.20%  
Best Local Similarity: 29.06%  
Query Match: 19.61%  
Length: 4577  
Matches: 451  
Conservative: 266  
Mismatches: 551  
Indels: 284

DB: 4 Gaps: 59  
 US-10-020-095-4 (1-1428) x US-09-241-606-1 (1-4577)  
 QY 12 LeuLeuCyValCysThrAlaAlaLeuAlaValAlaProGlyProCysArgPheLeuValThr 31  
 Db |||||  
 Db 86 CTCCTGCTCTCTCCGCCACAGACCGCTCAGCTCTGTGAAAAACCGCAATATATGCTCTG 145  
 QY 32 AlaProGlyIleIleArgProGlyGlyValAlaThrIleGlyValGluLeuGluHis 51  
 Db |||||  
 Db 146 GTCCCTCCCTGCTCCACACT---GAGACCACTGAGAGAGCGCTGTGCTCTCTAGAGCTAC 202  
 QY 52 CysProSerGlnValThrValIleAlaGluLeuLeuGlyThrAlaSerAlaLeuThrVal 71  
 Db |||||  
 Db 203 CTGAATGACAGACCTGTAAGTGCCTTCTTGAGTCTGACGGGAAACAGAGCGCTC 262  
 QY 72 ---SerValLeuGluAlaGluGly---ValPheGluGlySerPheLeuThrLeuThr 89  
 Db |||||  
 Db 263 TTCACTGACCTGAGCGGAGAAATGACGTAAGCTGCTGCTGCTC--- 310  
 QY 90 LeuProSerLeuProLeuAlaSerAlaAspGlu-----IleTyrgluLeuArgVal 106  
 Db |||||  
 Db 311 -----GCTGTCCTCCAAAGTCTTCAATCAATGAGAGTAATGCTTCACTGCTCAAGT 364  
 QY 107 ThrGlyArgThrGlnAspGluIleLeuPheSerAlaSerThrArgLeuSerPheGluThr 126  
 Db |||||  
 Db 365 AAAGGACCAACCAAGAA-----TTAAGAAAGCGGACCAACAGTATGTTAAGAAC 415  
 QY 127 LysAlaGlySerValPheIleGlnThrAspIleAlaLeuTyrglyProGlyGlnGluVal 146  
 Db |||||  
 Db 416 GAGGACACTGCTGCTTGTGTCACAGACAAATCAATCTCAAAACGAGGACAGACAGTG 475  
 QY 147 LysPheArgIleValIleThrLeuPheSerAspPheLeuProTyrglyThrSerLeuAla 166  
 Db |||||  
 Db 476 AAATTCCTGCTGCTCATGATGAAACCTTCAACCCCTGATGATGATTCATTCCTCACTA 535  
 QY 167 Leu---IleLysAspProLysSerAlaLeuIleGlnGlnThrLeuSerGlnGlnSerAsp 185  
 Db |||||  
 Db 536 GTATATCATTCAGATCCCAAGAAATCGCATCGACAAATGAGGAGTTTCCAGTTAGAG 595  
 QY 186 LeuGlyValIleSerLysThrPheGlnLeuSerSerHisAspIleLeuGlyAspTyrSer 205  
 Db |||||  
 Db 596 GGTGCTCCCAAGCAATTTCTTTTCCCTCTCATCAGACCTTCCAGGCTCTCTACAG 655  
 QY 206 IleGlnIleGlnValAlaAsp-----GlnThrTyrglyGlnSerPheGlnValSerGlu 223  
 Db |||||  
 Db 656 GTGCTGTCACGAAGAATTCAGGTGAGAGGACAGACAC---CCTTTCACCTGAGAGAA 712  
 QY 224 TyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrglySerMetAlaSer 243  
 Db |||||  
 Db 713 TTGTCTTCCCAAGTTGAAGTACAGATACAGTGCACAAAGATATATCACCATCTTGGA 772  
 QY 244 LysHisIleLeuAlaGlyThrIleThrAlaLysTyrglyTyrglyProValIleGlyAsp 263  
 Db |||||  
 Db 773 GAAAGAGATGATGATGATGTGCTGCTATACATATGAGAAAGCTGCTCCAGACAT 832  
 QY 264 ValIleThrLeuThrPheLeuProLeu----- 271  
 Db |||||  
 Db 833 GTGACTGTGACATTTGCAAGAAATATGATGACGCTTCCACTCCACGCTGAAAGATTCA 892  
 QY 272 ----- 276  
 Db 893 CAGGCTTTCTGTGAGAAATTCAGTGAACAGCTAAACAGCCATGCTGTTTATCAGAA 952  
 QY 277 LysIleAlaGlnIleThrLysThrPheLysIleAlaGlySerAlaAlaPheSerPheAsnAsp 296  
 Db |||||  
 Db 953 GTAAAA-----ACCAAGCTCTTCCAGCTGAAGAG-----AAGAG 988  
 QY 297 GluGluMetLysAlaAlaValMetAspSerSerAlaGlyLeuSerGluTyrglyLeuAspLeuSer 316  
 Db |||||  
 Db 989 TATGAATGAAAA-----CTTCACTGAGGCGCCAGATCCAA 1024  
 QY 317 SerProGlyProValGluIleLeuThr-----ThrValThrGluSerVal 331

Db 1025 GAAGAGGAACAGCTGTGGAATGACTGGAAGGACAGTCCAGTGAATTCACAGAACCATTA 1084  
 QY 332 ThrGlyIleSerArgAlaValSerThrAlaValPhePheGlyGlnHisAspTyrglyIle 351  
 Db |||||  
 Db 1085 ACCAACTTCA---TTTGAAAGTGAGCTCACACTTGTGACAGGGA-----ATT 1132  
 QY 352 GluPhePheAspTyrglyThrThrValLeuLysProSerLeuAlaPheThrAlaThrValLys 371  
 Db |||||  
 Db 1133 CCCTTCTTGG-----CAG 1147  
 QY 372 ValThrArgAlaAspGlyAlaGlnLeuThrLeuGluGluArgArgAlaAlaValIle 391  
 Db |||||  
 Db 1148 GTGGCTCTAGTAATGAGAAAGCGCTCCATACCA-----ATAAAGTCATA 1195  
 QY 392 ThrValThrGlnArgAlaValTyrglyTyrglySerAlaSerGlyAlaGlnLys 411  
 Db |||||  
 Db 1196 TTCAATC---AGAGAAATGAGCAAACTATTAC-----TCAATGCTACACGAGATGAG 1246  
 QY 412 MetGluAlaValGln---LysIleAlaValThr-----ValProGlnSerGlyThrPhe 428  
 Db |||||  
 Db 1247 CAGGCTCTGTACAGTTCTCTATCAACACCAAGTTATGGGTACCTCTTACTGTT 1306  
 QY 429 LysIleGluPhePro-----IleLeuGluAspSer 438  
 Db |||||  
 Db 1307 AGGCTCAATTAACAAGATCGTACCTGCTTACGCTACCAAGTGGGTGTCAGAAAGAAC 1366  
 QY 439 SerGluLeuGlnLeuLysAlaTyrglyPheLeuGlySerLysSerSerMetAlaHisSer 458  
 Db |||||  
 Db 1367 GAAGAGCACATCACTGCTTAT----- 1390  
 QY 459 LeuPheLysSerProSerLysThrTyrglyIleGlnLeuLysThrArgAspGluAlaIleLys 478  
 Db |||||  
 Db 1391 CTGTGCTTCCCAAGAGAGCTTGTGCTGACCTGAGCCATGCTCATGAACTAACCC 1450  
 QY 479 ValGlySer-----ProPheGluLeuValValSerGlyAlaLys-----Arg 492  
 Db |||||  
 Db 1451 TGTGGCATACTGACAGACAGTCCAGGACCAATTAATTCGATGATGAGGACACCTGCTGGG 1510  
 QY 493 LeuLysGluLeuSer-----TyrMetValValSerArgGlyGlnLeuValAlaValGly 510  
 Db |||||  
 Db 1511 CTGAAGAGCTCTCCCTTTATTTATGATATGACCAAGGAGGACATTTGCCAATCGGG 1570  
 QY 511 -----LysGlnAlaSerThr-----MetPheSerLeuThr 520  
 Db |||||  
 Db 1571 ACTCATGAGCTGCTGTGGAAGGACGAAGACATGAAGGCCATTTTTCATCTCAATCCCT 1630  
 QY 521 ProGluAlaSerThrProLysAlaCysValIleValTyrglyIleGluAspAspGly 540  
 Db |||||  
 Db 1631 GTGAAGTCAGACATTTGCTGCTGCTGCTGCTCATATGATGCTTTTACCTACCGGG 1690  
 QY 541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAlaLysIleLys 560  
 Db |||||  
 Db 1691 GACGTATTTGGGATTTCTGCAAAATATGATTTGAAAATTTGCTGGCCAAACAAAGTGAGT 1750  
 QY 561 LeuTyrglySerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580  
 Db |||||  
 Db 1751 TTGAGCTTACGCCCATCACAAGTCTCCAGCTCCACAGCCCACTGCGAGTCACAGCG 1810  
 QY 581 ThrGlnProAspSerIleValGlyIleValAlaValAlaAspLysSerValAlaLeuMetAla 600  
 Db |||||  
 Db 1811 GCT---CTTCAGTCCGCTGCGGCTCCGCTGCTGAGACCAAGCGTGCCTCATGAAG 1867  
 QY 601 AlaSerAlaAspIleThrMetGluAlaValHisGluLeu---GluLeuTyrglyThr 619  
 Db |||||  
 Db 1868 CTGATGCTGAGCTTCGGGCTTCGCTTACCACTGCTCAAGAAAGACCTCCTACT 1927  
 QY 620 GlyTyr-----TyrLeuGlyMet 625  
 Db |||||  
 Db 1928 GCGTTCCTGGGCTTTGAATGACGAGACGATGAAGATGTCATCAATGCTAATATGTC 1987  
 QY 626 PheMetAlaSer-----PheAla 631  
 Db |||||





QY	1276	AspleuAen---HisValAspleuAenValCysThrSerPheSerGlyProGly-----	1292
Db	4103	GAACCCAAAGCCCAACACAGCTCCAAATCTCCCAATGTGTCAATTAACAAGGAGCCGC	4163
QY	1293	---ArgSerGlyMetAlaLeuMetGluValAsnLeuLeuSerGlyPheMetValProSer	1311
Db	4163	TCTGCCTCCAAACATGGCGCATGGATGGATGTGAAGATGGCTCTGGCTTCAATCCCTGAAG	4222
QY	1312	GluAlaIleSerLeu-----SerGluThrValIlyIysValGluTyrAspHisGly	1328
Db	4223	CCAAACAGTAAATATGCTTGAAGAAGTCTAACCATGTGAACCCGACAGAGTCAAGACCAAC	4282
QY	1329	LysLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCysValAsnIleProAla	1348
Db	4283	CATGCTTGATTTACCTGTGATGAAGGTGTGCAAATCAGACATCAGAGCTTGTTCTTCAACGGTT	4342
QY	1349	ValArgAsnPheIysValSerAsnThrGlnAspAlaSerValSerIleValAspTyrTyr	1368
Db	4343	CTGCAGATGTCCCGATGAAGATCTCAAAACAGCCATATGTAAGAAAGTCAATGATTACTAC	4402
QY	1369	GluProArgArgGlnAlaValArgSerTyrAsnSer	1380
Db	4403	GAGACGATGATGTTGCATGCTGCTGATCAATGCT	4438

Pred. No.:	1,46e-132	length:	4079
Score:	1346.50	Matches:	419
Percent Similarity:	46.66%	Conservative:	238
Best Local Similarity:	29.76%	Mismatches:	488
Query Match:	18.32%	Indels:	263
DB:	4	Gaps:	54

QY 297 GUGUIMetLysAsnValMetAspSerSerAsnGlyLeuSerGlyIuThyLeuAspLeuSer 316  
 DB |||||  
 DB 989 TATGAATGAAA-----CTTCACACTGAGGCCGACATCCA 1024  
 QY 317 SerProGlyProValGluIleuThr-----ThyValThrGluSerVal 331  
 DB |||||  
 DB 1025 GAAGAAGAACAGTGTGGATTGACTGGAGAAGCACTCCAGTAATATCAAGAACATA 1084  
 QY 332 ThrGlyIleSerArgAsnValSerThrAsnValPhePheGlnHisAspTyrIleIle 351  
 DB |||||  
 DB 1085 ACCAAACTCTCA---TTGTGAAGAGTCACTCACTTTCAGACAGGA-----ATT 1132  
 QY 352 GluPhePheAspTyrThrThrValLeuLysProSerLeuAsnPheThrValThyValLys 371  
 DB |||||  
 DB 1133 CCTCTCTTGGG-----CAG 1147  
 QY 372 ValThrArgAlaAspGlyAsnGlnLeuThrLeuGluGluArgArgAsnValValIle 391  
 DB |||||  
 DB 1148 GTGGCCCTGTAGATGAGAAAGCGCTCCATATCCA-----ATAAAGTCATA 1195  
 QY 392 ThrValThrArgAsnTyrThrGlyIuTyrPheSerGlySerAsnSerGlyAsnGlnLys 411  
 DB |||||  
 DB 1196 TTCATC---AGAGGAATAGACAAACTTTAC-----TCCATCTTACACAGATGAG 1246  
 QY 412 MetGluAlaValGln---LysIleAsnTyrThr-----ValProGlnSerGlyThrPhe 428  
 DB |||||  
 DB 1247 CATGGCCTGTACAGTTCTCTATCAACACCAACGATTAGGGTACCTCTTACTGTT 1306  
 QY 429 LysIleGluPhePro-----LleuGluLysPheSer 438  
 DB |||||  
 DB 1307 AGGGTCATATTCACAGAGCTGATGCTGTTACCGCTTACAGTGGTGTCAAGAACAC 1366  
 QY 439 SerGluLeuGlnLeuLysAlaTyrPheLeuGlySerLysSerMetAlaValHisSer 458  
 DB |||||  
 DB 1367 GAAGAGCACATCACACTGCTTAT----- 1390  
 QY 459 LeuPheLysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLys 478  
 DB |||||  
 DB 1391 CTGTGTCTCTCCCAAGCAAGACTTGTTCACCTTGAAGCCCATGTCATGACATCACCC 1450  
 QY 479 ValGlySer-----ProPheGluLeuValValSerGlyAsnLys-----Arg 492  
 DB |||||  
 DB 1451 TGTGGCATACTCAGACAGTCCAGCAGCATATATCTGAATGAGAGCACCTGTCTGGG 1510  
 QY 493 LeuLysGluLeuSer-----TyrMetValValSerArgGlyGlnLeuValAlaValGly 510  
 DB |||||  
 DB 1511 CTGAAGACCTCTCTTTATATATCTGATATATGGCAAGAGGACATGTGTCCGAAGCTGG 1570  
 QY 511 -----LysGlnAsnSerThr-----MetPheSerLeuThr 520  
 DB |||||  
 DB 1571 ACTCATGACAGTCTGTGAAGCAGGAACATGAAGGCCATTTTCCATCTCAATCCCT 1630  
 QY 521 ProGluAsnSerThrThrProLysAlaCysValIleValTyrTyrIleGluAspGly 540  
 DB |||||  
 DB 1631 GTGAAGTCAGACATGCTCTGCTGCTGCTGCTCATCATGCTGTATTACTTACCGGG 1690  
 QY 541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560  
 DB |||||  
 DB 1691 GACGATTTGGGGATTCTGCAGAAATATGATGTGAATAATTGCTGGCCCAACAAGGAT 1750  
 QY 561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580  
 DB |||||  
 DB 1751 TTGAGCTTACGCCATTCACAAAGTCTCCAGCTCAGCCCACTGCGATGATCAGAGG 1810  
 QY 581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600  
 DB |||||  
 DB 1811 GCT---CTCAAGTCCCTGCGCCCTCCGCTGTGACCAAGCGCTGCTCATGAAG 1867  
 QY 601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeu---GluLeuTyrAsnThr 619  
 DB |||||  
 DB 1868 CCTGATGCTGAGCTCTGGCGTCTCTGGTTTACACCTGCTACCAAGAACAGCTCACT 1927  
 QY 620 GlyTyr-----TyrLeuGlyMet 625

DB 1928 GAGTCCCTGGGCGCTTGAATGACAGGACGATGAAGACTGCATCAATCTCATATATGTC 1987  
 QY 626 PheMetAsnSer-----PheAla 631  
 DB |||||  
 DB 1988 TATATTAAATGGAATCATATATATCTCCAGTATGACATCAATGAAGATATATGACG 2047  
 QY 632 ValPheGlnGluCysGlyLeuThrValLeuThrAspAlaAsnLeuThrLysAspTyrIle 651  
 DB |||||  
 DB 2048 TTCCTAGAGACATGGGCTTTAAAGGCATTCACAACTCAAAAGATTCTGTAACCCAAATG 2107  
 QY 652 AspGlyValTyrAspAsnAlaGlu-----Tyr 660  
 DB |||||  
 DB 2108 TGTCCACAGCTTCAACAGTATGAATGATGATGAGCTGAGGCTTACGTGATGTTTAT 2167  
 QY 661 AlaGluAsnPheMetGluGlnAsnGluLysIleValAspIleHisAspPheSerLeu 680  
 DB |||||  
 DB 2168 GAGTCAGATGTATGAGGAAGGCCAGTGCAGCGCTGGTGCATGTGAAGAG----- 2218  
 QY 681 GlySerSerProHis-----ValArgLysHisPheProGluThrTyrIleThrLeu 697  
 DB |||||  
 DB 2219 -----CTCACAGGAGACCGTACGAAGATACCTTCCCTGAGACATGATCTGGAT 2269  
 QY 698 AspThrAsnMetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIle 717  
 DB |||||  
 DB 2270 TTGGTGTGTGAATCTCAGCAGGGGTGAGTGAAGATGAACAGTCCGACACCATC 2329  
 QY 718 ThrSerThrValAlaThrGlyPheValIleSerGluAspLeuGlyLeuThrThr 737  
 DB |||||  
 DB 2330 ACCGATGAAAGGACGGGCGCTTGTGCTGTGAAGTGTGATGATGATCTTCC 2389  
 QY 738 ThrProValGluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSer 757  
 DB |||||  
 DB 2390 ACT---GCCTCTCCAGCTTCCAGGCTTCTTTTGAGAGCTTACATATGCTTACTCT 2446  
 QY 758 ValIleArgGlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAla 777  
 DB |||||  
 DB 2447 GTGATTCGTGAGAGGCTTACACTCAAGGCCAGCGCTTAACCTTCCCAATATGC 2506  
 QY 778 ThrGluValLysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSer 797  
 DB |||||  
 DB 2507 ATCCGGTCACTGTGCACATGGAAGCCTCCCGCTTCTCTGTCGACAGTGAAG 2566  
 QY 798 Glu-----IleAsnAlaThrGlnHisGlnGlnIleThrLeuValProSer 812  
 DB |||||  
 DB 2567 GAACAAGCGCTCAGTATGTGCAACGGCGGCA----- 2605  
 QY 813 GluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGlyGluIleProIle 832  
 DB |||||  
 DB 2606 -----ACTGTCTGGGAGATTAACCCCAAGTCAATTAGAAATGTGAATTTC 2653  
 QY 833 ThrValThrAlaLeuSer-----ProThrAla 841  
 DB |||||  
 DB 2654 ACTGTGACGAGAGGCACTAGACTTCAAGCTGCTGTGTGTAACCTGAAGAGACTA 2713  
 QY 842 -----SerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIle 856  
 DB |||||  
 DB 2714 CCTGAACGCGAAGAAAGACATCAAGCTCAAGCTCTGTGTGTAACCTGAAGAGACTA 2773  
 QY 857 GluLysSerTyrSerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThr 876  
 DB |||||  
 DB 2774 GAGAAAGAAACAACATTCATCCCTACTTTGTGCATCAGGTGGAGATTCT----- 2827  
 QY 877 LeuLysThrLeuSerPheSerPheProProAsnThrValThrGlySerGluArgValGln 896  
 DB |||||  
 DB 2828 ---GAAGATATATCCCTGAATCTGCACCAATGTGTGAAGATGTGCCCGCTTCT 2884  
 QY 897 IleThrAlaIleGlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArg 916  
 DB |||||  
 DB 2885 GTCTAGTTTGGGAGACATATATAGCTCTGCTCCATGCAAAACACACAAATCTTCTCAG 2944  
 QY 917 MetProTyrGlyCysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeu 936  
 DB |||||

```

Db      2945 ATGCCCTATGCGCTGGAGACAGAAATATGGTCTCTTTGGTCTCTTAACATCTATGTACTG 3004
Qy      937 AAPTGTLeuThrlYLeuYLeuGlnLeuThrlAspLeuLeuYLeuGlnYAlaLeuSerPhe 956
Db      3005 GATTATCTTAATGAACAACAGCAGCTTACTCCAGAGCTCAAGTCCCAAGGCCATTGGCGTAT 3064
Qy      957 MecArgGlnGlyTyrGlnArgGlnLeuLeuTyrGlnArgGlnArgGlySerPheSerAla 976
Db      3065 CTCAACTAGGCTTACAGAGACAGTTGTAACAAACACTATGATGGCTCTTACAGCAGCC 3124
Qy      977 PheGly---AspTyrAspProSer---GlySerThrlPheSerAlaPheValLeuArg 994
Db      3125 TTGGGGAGCCATATGAGAGAACCAAGGCAACACTGGCTCTACAGCTTTGTTCTTGAAG 3184
Qy      995 CysPheLeuGlnAlaAspProTyrTlleAspTlleArgGlnAsnValLeuLeuArgTyr 1014
Db      3185 ACTTTGGCCCAAGCTCGAGCTTACTTTCATTCATGAAGACACATTCACCAAGCCCTC 3244
Qy      1015 ThrTrpLeuYLeuYLeuGlnYLeuSerArgGlnYLeuPheTrpAspProGlyArgValIle 1034
Db      3245 ATATGGCTCTCCAGAGGACAGAAAGCAATGGCTGTTCAGAGCTGTGGTCACTGCTC 3304
Qy      1035 HisSerGlnLeuGlnGlnGlnYLeuYLeuSerProValThrlLeuThrlAlaTyrIleValThr 1054
Db      3305 AACATGCCATAAAGAGAGAGTAGAAGATGAAGTCACTCTCCGCTATATACCATC 3364
Qy      1055 SerLeuLeuGlyTyrArgYLeuTyrGlnProAsnTlleAspValGlnGlnSerTlleAspHe 1074
Db      3365 GCCCTTCTGGAGATTCCTCTCACAGTCACTCACTGTGTCGCCAATGCCCTGTTTGC 3424
Qy      1075 LeuGlnSer-----GlnPheSerArgGlyIleSerAspAsnTyrThr 1088
Db      3425 CTGGAGTCAAGCTGGAGAGACAGACAAAGAGGGAGCCATGGC---AGCCATGATATATACC 3481
Qy      1089 LeuAlaLeuTleThrlYAlaLeuSerSerValGlySer---ProYValYLeuGlnAla 1107
Db      3482 AAAGACACTGGCGCTATGCTTTTGGCTCTGGCAGGTAACAGCAACAAAGAGAGAGTA 3541
Qy      1108 LeuAsnMetLeuThrlTrpArgAlaGlnGlnGlnGlyLeuTyrGlnPheTrp----- 1124
Db      3542 CTCAGTCACTTAATGAGAAAGCTGTGAAGAAAGCAACTGTTCATTCAGGAGGCCCT 3601
Qy      1125 ValSerSerGlnSerYLeuSerAspSerTrpGlnProArg-----SerLeuAspIle 1142
Db      3602 CAGAAACCCAGGACAGCAGTGGGCGATTTTACGAACCCAGGCTCCCTGTGAGAGTG 3661
Qy      1143 GlnValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGln-----ThlSer 1159
Db      3662 GAGATGACATCTTATGTGCTCTCTGCT---TATCTCAGGCGCCAGCCAGCCCAACCTCG 3718
Qy      1160 GlnGlyIlePro-----TlleMetArgTrpLeuSerArgGlnArgAsnSerLeu 1175
Db      3719 GAGGACCTGACCTCGCAACCAACATCGTGAAGTGAATGATCAAGAAAGCAGCAAAATGCCAG 3778
Qy      1176 GlyGlyPheAlaSerThrlGlnAspThrlYAlaLeuYValaLeuSerGlnPheAla 1195
Db      3779 GCGGCTTCTCTCTCCACCAAGACACAGTGTGGCTCCATGTCTGTCTCAAAATATGGA 3838
Qy      1196 AlaLeuMetAsnThrlGlnArgThrlAsnTlleGlnValThrlValThrlGlyProSerSerPro 1215
Db      3839 GCCCCACATTTACC---AGGACTGGGAAGGCTGACAGTGAATTCACATTCAGTCTTCAAGGG 3895
Qy      1216 SerProLeuAlaValaGln----- 1222
Db      3896 ACATTTTCCAGCAATTCCAAGTGGACAACAACATCGCTGTACTGCAAGCAGGTCTCA 3955
Qy      1223 -----PrometAlaValaAsnTlleSerAlaAsnGlyPheGlyPheAlaIleCys 1238
Db      3956 TTGGCAAGAGCTGCTGGGGATACAGCATGAAGAGAGAGAGATGTGTCTACTCTC 4015
Qy      1239 GlnLeuAsnValaValTyrAsnVal 1246
Db      4016 CAGACCTCTCTTGAATAATCAATATT 4039

```

```

RESULT 3
US-08-447-411-1
; Sequence 1, Application US/08447411
; Patent No. 5773243
;
GENERAL INFORMATION:
; APPLICANT: FRITZINGER, DAVID C.
; APPLICANT: BREDEHORST, REINHARD
; APPLICANT: VOGEL, CARL-WILHELM
; TITLE OF INVENTION: DNA ENCODING COBRA C3, CVF1, AND CVF2
; NUMBER OF SEQUENCES: 81
;
CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLOM, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
;
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
;
CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/447,411
; FILING DATE:
; CLASSIFICATION: 435
;
PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/043,747
; FILING DATE: 07-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Obloin, No. 5773243man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 1126-101-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
;
SEQUENCE CHARACTERISTICS:
; INFORMATION FOR SEQ ID NO: 1:
; LENGTH: 5211 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
;
FEATURE:
; NAME/KEY: CDS
; LOCATION: 9..4961
;
US-08-447-411-1
;
Alignment Scores:
Pred. No.: 1,376-92 Length: 5211
Score: 973.50 Matches: 392
Percent Similarity: 40.56% Conservative: 280
Best Local Similarity: 23.66% Mismatches: 650
Query Match: 13.25% Indels: 335
DB: 1 Gaps: 62
US-10-020-095-4 (1-1428) x US-08-447-411-1 (1-5211)
Qy      1 MetGlnGlyProProLeuLeuThrlAlaAlaHisLeuLeuCysValaCysThrlAlaLeu 20
Db      9 ATGAGGGGATGGCTCTTATCTGCGTGGCTGTATGATTTGCTTTCCAGGGTCT--- 65
Qy      21 AlavaAlaProGlyProArgPheLeuValThrlAlaProGlyTlleIleArgProGlyGly 40
Db      66 -----TCCACGGGGCTCTATACCCCTCATGACCCCTGCTGTTTGGCAACAGACACA 119
Qy      41 AsnValThrlIleGlyValGlnLeuLeu---GlnHisCysProSerGlnValThrlValIle 59
Db      120 GAAGAGCAATTTTGGTGGAGGCCCATGAGACAGTACTCAAAATCGCTTGAACATCTTT 179

```

QY 60 Ala-----GluLeuLeuysThrAlaSerAsnLeuThrValSerValLeuGlu 75  
 Db 180 GTTCATGATTTTCCAGGAGGAGGAAACCTTGTTCGAAGCAGATGATGATCAATCAG 239  
 QY 76 AlaGluGlyValPheGluIuysGlySerPheIysThrLeuThrLeuProSerLeuProLeu 95  
 Db 240 GCAGGAGACATGTTGTCTACT-----CCAACTATAAGGTTCTCTCAAAAGAACTG 290  
 QY 96 AsnSerAlaAspGluIleTyrgluLeuArgValThrGlyArgThrGlnAspGluIleLeu 115  
 Db 291 AATAAGACCTCCAGGAAATATCATGATGCTTGGAAGTAACTGCTCTCAAGTGGCA 350  
 QY 116 PheSerAsnSerThrArgLeuSerPheGluThrLysArgIleSerValPheIleGlnThr 135  
 Db 351 TTGAAAAGGTGGTTCTCTTCTTCAACAGAGTGC-----TTTGTTGTTCACTCCAGACA 404  
 QY 136 AspLysAlaLeuTyrglyProIysGlnGluValLysPheArgIleValThrLeuPheSer 155  
 Db 405 GATTAAGGATCTATACACCAAGCTCTCCAGTGCCTATCGTCTTTCTGTG----- 458  
 QY 156 AsppheLysProTyrglySerLeuAsnIleLeuIleLysAspProLysSerAsnLeu 175  
 Db 459 GATCAACACATGCACGAGATGCACAAACTGTGATTTGTCAG----- 500  
 QY 176 IleGlnGlnTrpLeuSerGlnIleSerAspLeuGlyValIleSerLysThrPheGlnLeu 195  
 Db 501 -----TTTCAGACTCCAGAAAGCATTTGT-----GTC 527  
 QY 196 SerSerHisProIle----- 200  
 Db 528 AGCTTCAACACAGTCAATCCATCAGGCTGATCCGGCTTACATTTACAGAGCTTGC 587  
 QY 201 ---LeuGlyAspTrpSerIleGlnValGlnValAsn-----AspGlnThrTyrgly 216  
 Db 588 AGTTTGGGACATGCAGAGCTGTGGCCAAATATGAACATTCACACGAAAGAGCTACCT 647  
 QY 217 GlnSerPheGlnValSerGluTyrglyValLeuProLysPheGluValThrLeuGlnThrPro 236  
 Db 648 GCATATTTTGTGATCGAGATATATGTTTACCAAGCTTTGAAAGTCCGTGTGCAACCATCA 707  
 QY 237 -----LeuTyrglySerMetAsnSerLysHisLeuAsnGlyThrIleThrAlaLys 253  
 Db 708 GATTAAGTTCTTTACATGATGAGAT---AAAAATTTCACAGTGTCTATCACTGCAAG 764  
 QY 254 TyrglyTyrglyLeuProValLysGlyAspValThrLeuThrPheLeuProLeuSerPhe 273  
 Db 765 TACTTATATGGAAGAAAGTGAAGGT---GTGGCTTTGTGCTCTTGGAGTCAAAATA 821  
 QY 274 TrpGlyLysLysAsnIle-----ThrLysThrPheLysIleAsnGlySer 289  
 Db 822 GATGATGCTAAAGATATTCAGACTCAGTCAAGAGAAATTCGATTTTATGAGAGAT 881  
 QY 290 AlaAsnPheSerPheAsnAspGluGluMetLysAsnValMetAspSerSerAsnGlyLeu 309  
 Db 882 GGGGAGCAACACTAAAGAGATACATACGTTCCGATTTCAAGATCTCATCAGCTT 941  
 QY 310 SerGluTyrglyLeuAspLeuSerProGlyProValGluIleLeuThrThrValThrGlu 329  
 Db 942 GTTGTCTACTCTGTATGATCT-----GTAACTGATGATACAGAA 983  
 QY 330 SerValThrGlyIleSerArgAsnValSerThrAsnValPhePheLysGlnHisAspTyrgly 349  
 Db 984 TCAGGACGATATGATGATGATGATGAGCAGCGGATTCATATTTGATGATCTCCCTAT 1043  
 QY 350 IleIleGluPhePheAspTyrglyThrThrValLeuLysProSerLeuAsnThrAlaThr 369  
 Db 1044 CAGATCTACTCTCAAAAACCCCAAAATTTTCAAGCCAGAGATGCAATATGAACAGC 1103  
 QY 370 ValLysValThrArgAlaAspGlyAsnGlnLeuThrLeuGluGluArgArgAsnVal 389  
 Db 1104 GTGTATGTTTACCAACCTGATGGCTCAGCAGCTGCC-----CAT 1142  
 QY 390 ValIleThrValThrGlnArgAsnTyrglyTyrglySerGlySerAsnSerGlyAsn 409

Db 1143 GTGCCAGTGATATCAGAGCCATTCATCTGAG-----GGAACC 1181  
 QY 410 GlnLysMetGluAlaValGlnLysIleAsnTyrglyThrValPro-----GlnSerGly 426  
 Db 1182 ACTTTGAGTGAGGAGACTCTTAAGCTCATTTGAAACACACCACTGAACATTTAAAGC--- 1238  
 QY 427 ThrPheLysIleGluPheProIleLeuGluAspSerSerGluLeuGlnLeuLysAlaTyrgly 446  
 Db 1239 -----CTACCATCATCTGTATGAACTTAACTAACATGAGAGACTCCCAAGAGA 1283  
 QY 447 PheLeuGlySerLysSerSerMetAlaValHisSerLeuPheLysSerProSerLysThr 466  
 Db 1284 CGCCAGGCAATTAAGTCCATGACAGCCACAGCTTACCAACCCAGGAGAGGCTGTAAAC 1343  
 QY 467 TyrglyGlnLeuLysThrArgAspGluAsnIleLysValGlySer-----ProheGlu 484  
 Db 1344 TATCTTCATGTATGCCATTATCATCTACAGAGATTAAGCCCGAGATTAACCTTACTGTCAAT 1403  
 QY 485 LeuValIleSerGlyAsn-----LysArgLeuLysGluLeuSerTyrglyMetVal 500  
 Db 1404 TTCAATGTGAGGCGCAATTCGAATTCAGTGAACCGATCAATATTTACATCTCTCAT 1463  
 QY 501 ValSerArgGlyGlnLeuValAlaValGlyLys-----GlnAsnSer 514  
 Db 1464 TTGAATTAAGGAGAAATTTTCAAGGTTGAGGAGCAACCCAGAGAGATGGCAGATCTG 1523  
 QY 515 ThrMetPheSerLeuThrProGluAsnSerTrpThrProLysAlaCysValIleValTyrgly 534  
 Db 1534 GTGACCATGAATCTGCATATCATCTCCAGATCTCATCTCTCTCCGTTGTGGCTTAC 1583  
 QY 535 TyrglyLeuAspAspGlyGluIleIleSerAspValLeuLysIleProValGln----- 552  
 Db 1584 TACCAAGTGGAAATTAAGAAATTTGCTGATTTCTGTGGTGAATGTAAGATACC 1643  
 QY 553 -----LeuValPheLys-----AsnLysIleLysLeuTyrgly 562  
 Db 1644 TGCATGGAGACGTTGTTGTGAAGAGCGTCTTCCAGACAGCATGCAATACAA----- 1697  
 QY 563 TrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerValThrGln 582  
 Db 1698 -----AAGCCAGAGCTCAAGCAAGAAATCAAAATGGA---GGGAT 1736  
 QY 583 ProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsnLysSer 602  
 Db 1737 CAGGTGCTCGGTTGTGCTTGTGCTGTGGCAAGAGATATGATTTCTCATATATAA 1796  
 QY 603 AsnAspIleThrMetGluAsnValAlaHisGluLeuGluLeuTyrglyAsnThrGlyTyrgly 622  
 Db 1797 TATAAGATTTAGCCAAAGCTAAAGATATGAGACACAAATGAAGAAAGAGTACTTGGCTGTACA 1856  
 QY 623 LeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTyrglyValLeuThr 642  
 Db 1857 GCTGGCAGTGGCCAGAAATATCTGGGTGTGTTGAAGATGCTGGACTGGCTTCAACAC 1916  
 QY 643 AspAlaAsnLeuThr----- 647  
 Db 1917 AGCACTAATCTACACCAAAACAGATCAGCTCAAAAGTGTCTTCAAGCTTCAAAATGG 1976  
 QY 647 ----- 647  
 Db 1977 AGCGCTCCAGTTCGTTTGTGCTGTGACAGCAAAAGCAAGAACCGGACAGTTTCAG 2036  
 QY 648 -----LysAspTyrglyLeuAspGlyValTyrglyAsn----- 657  
 Db 2037 GATCAAGCGCTCGTAATATGCTGTGAAGATGATGATGATGAGAACCCATGGGATACCT 2096  
 QY 658 AlaGluTyrglyAlaGluArgPheMetGluGluLysGly----- 670  
 Db 2097 TGTGAAAAGCGTGCAAAATACATCCAGAGGAGATGCTTTGAAGGCTGCTTCTTGA 2156  
 QY 671 -----HisIleVal----- 673

Db 2157 TGTGTACTACATCAAGGATCCGAGATGATAAACAAGGAGAGGAGCTTTCTG 2216  
 Qy 674 -----AspIleHisAspPheSerLeuGlySerSerProHisVal---ArgLeHis 689  
 Db 2217 GCAGAACTGATTTTGAAGATGAACCTTTGGAGATGACAAACATCATCTCCAGGCTTGAT 2276  
 Qy 690 PheProGluThrTrpIleTrpLeu-----AspThrAsnMetGly 702  
 Db 2277 TTTCCTCGAGATGGTGTGGCTAAACAGAAATTGACCGGGAGACCTTAACATCAAGAGG 2336  
 Qy 703 TyrArgIleTyrGlnGluPheGluValThrValPro-----AspSerIleThr 718  
 Db 2337 ATTTCA-----AGCAAGACAGTACCTTTTATCTGAGGAGATTCCATCA 2381  
 Qy 719 SerTrpValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThr 738  
 Db 2382 ACCTGG-----GAGTTGCTGGCTGGCTGGGCTTTTCAACCCACC 2417  
 Qy 739 -----ProValGluLeuGlnAlaPheGlnProPhePheIlePhe 751  
 Db 2418 AAGGAGATCTGTGGCTGAACCGATGATAAATACAGCATGAAGACTTCTTATTCAT 2477  
 Qy 752 LeuAsnLeuProTyrSerValIleArgGlyGluGluPheAlaLeuGluIleThrIlePhe 771  
 Db 2478 CTTGCAGCTGCATATTCAGTAGTAGAATGAGATGAGATGAGATTCGACTATTCTGATC 2537  
 Qy 772 AsnTyrLeuLysAspAlaThrGluValIleValIleIleGluLysSerAspLysPheAsp 791  
 Db 2538 AACCTGCTGACGAGATATTATATGTCGAGTGAACATCATATACAAACCCAGCTTC--- 2594  
 Qy 792 IleLeuMetThrSerSerGluIleAsnAlaThrGlyHisGlnGlnThrLeuLeuValPro 811  
 Db 2595 -----TCAGTGGCTTCCACAGAGACAAAGATACGACAGAGCTCCCATTTAA 2645  
 Qy 812 SerGluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGly-----Glu 829  
 Db 2646 GCCCTGCTCCAGACAGATCATTTGTGTAGTCCATTAGACAGAGATTCATCATGAT 2705  
 Qy 830 IleProIleThrValThrAlaLeuSerProThrAlaSerAspAlaValThrGluMetIle 849  
 Db 2706 GTTAGGATTATAGCAAGTGTCCGGGAGAGATGGCATCATGATGCTGAGAGAAAGAACTG 2765  
 Qy 850 LeuValLysAlaGluGlyIleGluLysSerThrSerGlnSerIleLeuLeuAspLeuThr 869  
 Db 2766 AAGGTTGACCTGAAGAGGGAACGAAATAATTGACTATATTATTAAGATGACCAAGT 2825  
 Qy 870 AspAsnAlaGluGlnSerThr-----LeuLysThrLeuSerPheSerPheProAsn 887  
 Db 2826 GTAAAGAGAGTGTGGAACCCAGGAACTAAACGTCATAGCCAAATTAATTAGATGACAA 2885  
 Qy 888 ThrValThrGlySerGluArgValGlnIleThrAlaIleGlyAsp-----Val 903  
 Db 2886 GTGCTGATATACGAAGTTGAGAACCGAGATTTCTGTGTAGGAGCCCTGGCTCAGATT 2945  
 Qy 904 LeuGlyProSerIleAsnGly-----LeuAlaSerLeuIleArgMetProTyrGlyCys 921  
 Db 2946 ATTGAACATCATTTGATGAAGTAATCAATCATCATCTATTATCTCTTGGCGCT 3005  
 Qy 922 GlyGluGlnAsnMetIleAsnPheAlaProAsnIle-----TyrIleLeuAspTyr 938  
 Db 3006 GGGGAGCAAAATATGATCATCATGATCATTCATTCGACCTACCTACTACTGAGACCA 3065  
 Qy 939 LeuThrLysLysGlnLeuThrAspAsnLeuLysGluValAlaLeuSerPheMetArg 958  
 Db 3066 ACAAGGCGTGGGAATCTTGTTGGATGCGAGACTGAAGCATATAAACAAGATCATG 3125  
 Qy 959 GlnGlyTyrGlnArgGluLeuLysTyrGlnArgGlySerPheSerAlaPheGly 978  
 Db 3126 ACTGTTATGCCAGACAGATGCTATCAAAAGAGATCATCTCTATGACGACATTTTCA 3185  
 Qy 979 AsnTyrAspProSerGlySerThrTrpLeuSerAlaPheValLeuArgCysPheLeuGlu 998  
 Db 3186 AAC-----CGTGATCTAGTTCTTGCTTAACAGCATATGTGTGAAGCTTAGCCATG 3239

Qy 999 AlaAspProTyrIle-----AspIleAspGlnAsnValLeuHisArgThrTrpTrpLeu 1017  
 Db 3240 GCTTCCAAACATGTAATAAGACATTTAGCCATGAATATTATTTGGAGGTGGAATGGCTC 3299  
 Qy 1018 ---LysGlyHisGlnLysSerAsnGlyGluPheThrAspProGlyArgValIleHisSer 1036  
 Db 3300 ATTCTGAACAGGACCAACACAGATGAGTGTCAAAAGAAATGCCCTGTGATCCATGGA 3359  
 Qy 1037 GluLeuGlnGlyValAsnLysSer-----ProValThrLeuThrAlaTyrIleVal 1053  
 Db 3360 GAATGCTGGAGAGAACTTAAGGCTGTAACCAAGACATCTTTACAGCATTCATTGTG 3419  
 Qy 1054 ThrSerLeuLeuGlyTyrArg-----LysTyrGlnProAsnIle-----Asp 1067  
 Db 3420 ACTGCATATTGGAATCCAGATCATGCTGCAAAAGAACAAATCATATTCTAGACAGCAGC 3479  
 Qy 1068 ValGlnGluSerIleHisPheLeuGluSerGluPheSerArgGlyIleSerAspAsnTyr 1087  
 Db 3480 ATCAATAGAGCCACAGATTTATTTACTCAAAAGATGAGAAA---CTGCAAGGCTTTAC 3536  
 Qy 1088 ThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGlySerProLysAlaLysGluAla 1107  
 Db 3537 ACTACAGCCCTCACAGCTTATGCTTGGCT-----GCTGCAAGCCGA 3578  
 Qy 1108 LeuAsnMetLeuThrTrpArgAlaGluGlnGluGlyGlyMetGlnPheTrpValSerSer 1127  
 Db 3579 CTCATATATGACAGGATCATCATGCGACATCAACAGAAAGAAATGCTTG----- 3629  
 Qy 1128 GluSerLysLeuSerAspSerTrpGlnProArgSerLeuAspIleGluValAlaAlaTyr 1147  
 Db 3630 -----GAAATATTAATGCTCGCACCATCATTAATTGAGGACACTTCCAT 3674  
 Qy 1148 AlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGluGlyIleProIleMetArgTrp 1167  
 Db 3675 GCCTGTGGCCCTGCTGAAATGAAGAAATTTGCTGAGTGCCTGTGATGACAGTGG 3734  
 Qy 1168 LeuSerArgGlnArgAsnSerLeuGlyGlyPheAlaSerThrGlnAspThrThrValAla 1187  
 Db 3735 CTGATGATGCAAAATTTATGGGGGAACATATGACAAACCCAGACAGATTAATGCTG 3794  
 Qy 1188 LeuLysAlaLeuSerGluPheAlaAlaLeuMetAsnThrGluArg---ThrAsnIleGln 1206  
 Db 3795 TTTCAGACTCTGCTGAATATGAGATTCAATCCATCCATCAGACATTAATTTTATGAT 3854  
 Qy 1207 ValThrValThrGlyProSerSerProSerPro-----LeuAla 1219  
 Db 3855 ATTTCTATTAACCTGCCAGAACGAAAGTACCTGAAGATTAATGATTAAGTAAT 3914  
 Qy 1220 ValValGlnProMetAlaVal-----AsnIleSerAlaAsn 1231  
 Db 3915 GCTGTCCAGCCCGGACAGTATGAGACCAACTCAGCAAGACCTTCACTGTCTCAGATCA 3974  
 Qy 1232 GlyPheGlyPheAlaIleCysGlnLeuAsnValValTyrAsnValLysAlaSerGlySer 1251  
 Db 3975 GGTGATGAAAGCAACAATGACCATTTGACGCTATATATATACACATATG----- 4025  
 Qy 1252 SerThrArgArgArgSerIleGlnAsnGlnGluAlaPheAspLeuAspValAlaVal--- 1270  
 Db 4026 ---AGGAGAGATGCAAATGTTTGCACAA---TTCCATCTTGATGTTTCTGTGAA 4076  
 Qy 1271 -----LysGluAsnLysAspAspLeuAsnHisValAlaAspLeuAsnVal 1284  
 Db 4077 AACGTGAATTTGAATTAACACGCAAGGAGGAGGCAAGCACCCTCAGAGCTTAATAATC 4136  
 Qy 1285 CysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluValAlaLeuLeu 1304  
 Db 4137 TGCACTATGATCTGGAGAAAGTTGATTTCAATATGACATTAATTAATTTCTATGCTG 4196  
 Qy 1305 SerGlyPheMetValProSerGluAlaIle-----SerLeuSerGluThr 1319  
 Db 4197 ACTGTTTTCCTGATGCTGAAGACCTTAAAGGCTTTCTAACGAGTGCAGACAGATAC 4256





QY	300	lysaenValMetAapSerSeranelyLeuserGluTyrLeuAapLeuserSerProGly	319
Db	1013	AAAGTCTTGTAGCTGCTCC-----	10333
QY	320	ProValGluIleLeuThrThrValThrGluSerValThrGlyIleSerArganValSer	339
Db	1034	-----ACCCTCATCTTGCACCTAGGACGACGTGACGTGTGCAGGACAGGCC	10787
QY	340	ThraenValPhePheIleGlnIleAapTyrIleIleGluPheAapTyrThrVal	359
Db	1079	AGCGGAGATCCCATGTCAGCTCTCCCTACACGATTCACCTTACCAAGACACCCAAATAC	1138
QY	360	LeuIlePProSerLeuAanPheThrIalathrValIysValThrArgAlaAapGlyAan---	378
Db	1139	TTCAAAACCAAGAAATCCCTTTACCTCATGCTTCGTGACCAACCTCGATGGCTCTCTA	1198
QY	379	-----GlnLeuThrLeuGluGluArgIArgAanAanValIleThrValIThrGlnArg	396
Db	1199	GCCTTACCAAGATCCCGGTGGCAGTCCAGGGCGAGGACATGTGCATCTCTTAACTCCAGGGA	1258
QY	397	AenTyrThrGluTyrTyrPserGlySerAanSerGlyAanGlnIysMetGluAlaValGln	416
Db	1259	GATGGCGGTGGCCAAACTCAGCATCAACACACCCACCCAGAG-----	1303
QY	417	lysIleAenTyrThrValProGlnSerGlyThrPheIysIleGluPheProIleLeuGlu	436
Db	1304	-----CCCTTGAGCATCAACGGTCCGACAGAAACAGAGAGCTCTCG	1345
QY	437	AapSerSerGluLeuGlnLeuIysAlaIyrPheLeuGlySerIysSerSerMetAlaVal	456
Db	1346	GAGGCAAGACCA-----GCTTACCAAGACCATGAGCGTCTG	1381
QY	457	HisSerLeuPheIysSerProSerIysThrTyrIleGlnLeuIysThr-----	472
Db	1382	CCCTACAGACACCGTGGGCACTCCAACATTAATCTCATCTCTCAGTCTCATGTCATACAG	1441
QY	473	-----ArgAapGluAanIleIysValGlySerProPheGluLeuValIalSerGlyAan	490
Db	1442	CTCAGACCCGGGGAGACCTCAACCTCTCTCTGCGAATGACCGGCGCCACAG	1501
QY	491	lysAaIleuIysGluLeuSerTyrMetValValSerArgIleGlnLeuValAlaValGly	510
Db	1502	GCCAAAGATCCGCTACTACACCTCATCATCAAGAACAGGGCAGGCTTTGAAGCGGGA	1561
QY	511	lys-----GlnAanSerThrMetPheSerLeuThrProGluAanSer	524
Db	1562	CGCCAGGTCGAGAGCCCGCGCAGGACCTGGTGTGCTGCCCTTCCATCACCGAC	1621
QY	525	TyrThrProIysAlaCysValIleValIleValTyrTyr-----IleGluAapAapGly	540
Db	1622	TTTCATCCCTTCTCTCGGCTGGTGGCGTACACTACACGTCGATCGGCGCAGGGCGCAGAG	1681
QY	541	GluIleIleIleSerAapValleuIysIleProValGln-----LeuVal	554
Db	1682	GAGGTGTGGCGCACTCCGTGGGTGACGTCAAGACATCTCGCTGGCTGGCTGTG	1741
QY	555	PheIysAanIysIleIysLeuTyrTyrPserIysValIysAlaGlu-----	569
Db	1742	GTAATAAGCGGC-----CAGTCAAGAGACCGGACGCTGTA	1777
QY	570	ProSerGluIysValSerIleuArgIleSerValThrGlnProAapSerIleValGlyIle	589
Db	1778	CTTGGGACAGCAATACCTCTGAAGTAAAGGTATCACAGGGGCGCGGGGTGTG---CTG	1834
QY	590	ValAlaValAapIysSerValAanIleuMetAanIalSerAanAapIleThrMetGluAan	609
Db	1835	GTGGCGGTGACCAAGCGCGTGTCTGTCTGATAATAGAAACAACATGACCGACGATGAG	1894
QY	610	ValValIleIleGluLeuGluLeuTyrAanThrGlyTyrTyrLeuGlyMetPheMetAanSer	629
Db	1895	ATCTGGACGTGTGGAGAGGACAACTCGCTTCACCCCGGCGACGTGGAAAGATTAAC	1954

QY	630	PheAlaValAPheGlnGluCysGlyLeuThrValLeuThrAspAlaA-----	644
DB	1955	GCCGGTGTCTTCTCCGACGAGGGCTGACCTTACAGACGACGACGAGCCGACGACCGCC	2014
QY	644	-----	644
DB	2015	CAGAGGGGAGAACTTTCAGTGGCCGACGACCGCCGCGACCGCTTCGGTGCAGCTC	2077
QY	645	-----	651
DB	2075	ACGGAGAAAGCGAATGACAAAGTCGCGCAATGATACCCCAAGAGAGCTGGCAAGTCGCGAG	2134
QY	652	AspGlyValAlaYrAspAspN-----	666
DB	2135	GACGGCATGCGGAGAAACCCCATGAGTCTTCGTGCGACCGCCGAGACCCGCTTCAATCC	2199
QY	667	GluAaGlu-----	676
DB	2195	CTGGGCGAGGGCTGCAGAAAGTCTCTCTGAGACTGCTGCACATACAGAGCGTGGCG	2255
QY	677	-----	680
DB	2255	CGGACAGCACGGCGGGCCAGCCACCTGGGCGCTGGCCAGAGATACCTGATGAGACATC	2314
QY	681	GlySerSerProHisVal-----	696
DB	2315	ATTGCAGAAAGAAATCATCGTTCCCGAAGTACGATCCAGAGAGCTGGCTGTGAAACGTT	2374
QY	697	-----	706
DB	2375	GAGACATTGAAAGACCCACGAAATATGGAATCTCTACGAAGCTC-----	2413
QY	707	GlnGluPheGluValAlaThrAlaProAspSerIleThrSerThrValAlaThrGlyPheVal	726
DB	2420	-----	2472
QY	727	IleSerGluAspLeuGlyLeuGlyLeuThrThrThrProValGluLeuGlnAlaPheGln	746
DB	2474	ATGTCGCGAAGAAAGGAGATC-----	2533
QY	747	ProPhePheIlePheLeuAsnLeuProTyrSerValIleArgGlyGluGluPheAlaLeu	766
DB	2531	GACTTCTTCACTGCACCTCGCGCTACCTCTACTCTGTGTTCGAAACGACAGAGTGGAAATC	2599
QY	767	GluIleThrIlePheAsnTyrLeuIlyAspAlaThrGluValIlyValIleIleGlu---	785
DB	2591	CGAACCGGTCTCTACATATAC-----	2644
QY	786	--LysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsnAlaThrGlyHis	804
DB	2648	CTCCACCAATCAGAGCTTCTGAGCGCTGGCCACACCAAGAGG-----	2698
QY	805	GlnGlnThrIleLeuValProSerGluAspGlyAlaThrValLeuPheProIleArgPro	824
DB	2696	CAGCAGACCAATACCATCCGCCCAAGTCTCGTTGCTGCATATGTCATGTCGCG	2755
QY	825	-----	842
DB	2756	CTTAAAGACCGGCTTCGAGAGAGTGAAGTCAAGGCTGCTGTCTACATCATTTCACTACGT	2815
QY	843	AspAlaValAlaThrGlnMetIleLeuValIlyAsnGluGlyIleGluYrSerTyrSerGln	862
DB	2816	GAGGGTGCAGGAAGTCCCTGAAGTGTGTGCGGAGGAATTCAGAAATGACAAACACTGTG	2877
QY	863	SerIleLeuLeuAspLeuThrAspAsnArgLeu-----	880
DB	2876	GCTGTT-----	2923
QY	881	SerPheSerPheProProAsnThrVal-----	894
DB	2930	-----	2983
QY	895	-----	910



```

REFERENCE/DOCKET NUMBER: 102286.377
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 526-6000
TELEFAX: (617) 526-5000
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 5056 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-132-271-2

Alignment Scores:
Pred. No.: 4,44e-88 Length: 5056
Score: 931.00 Matches: 381
Percent Similarity: 40.88% Conservative: 287
Best Local Similarity: 23.32% Mismatches: 609
Query Match: 12.67% Indels: 358
DB: 3 Gaps: 63

US-10-020-095-4 (1-1428) x US-09-132-271-2 (1-5056)
QY 3 GlyProProLeuLeuThrAlaAlaHisLeuLeuCySValCysThrAlaAlaLeuAlaVal 22
   |||||
DB 65 GGTCCACGCTGCTG-----CTCCTGCTACTAACCCACCTCCCTGCTGCTG 112
QY 23 AlaProGlyProAlaGlyPheLeuValThrAlaProGlyIleIleArg----- 37
   |||||
DB 113 GGG---AGTCCCATGACTCTATCATCAACCCCAACATCTTGCGCGTGGAGAGCGAGAG 169
   |||||
QY 38 -----ProGlyGlyAsnValThrIleGlyValGluLeuLeu 49
   |||||
DB 170 ACAAATGCTGCTGAGAGCCGACGAGCGGAGGAGATGTTCCAGTCACTGTTACTGTCAC 229
   |||||
QY 50 GluHis-----CysProSerGlnValThrValIleValIleGluLeuLeu 63
   |||||
DB 230 GACTTCCCGACGCAAAAACTAGTGTCTGTCCAGTGGAGAGACTGTGTCGACCCCTGCCAC 289
   |||||
QY 64 LysThrIleAsnLeuThrValIleSerVal---LeuGluAlaGluGlyValPheGluLys 82
   |||||
DB 290 AACCATGCGGCAACGTACCTTCAACATCCACGACCAACAGGAGTTCACTGACGAAAG 349
   |||||
QY 83 GlySerPheLysThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGluIleTyr 102
   |||||
DB 350 GGGGCAACAAGTTCGTGACCGTGCAGGCC-----ACCTTGGGACCAAGTGTGTG 400
   |||||
QY 103 GluLeuArgValIleThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThrArgLeu 122
   |||||
DB 401 GAGAAAGTGTGCTGCTGTCAGCCTGCAGACCGCGTACCTC----- 439
   |||||
QY 123 SerPheGluThrLysArgIleSerValPheIleGlnThrAspLysAlaLeuTyrLysPro 142
   |||||
DB 440 -----TTCAATCCAGACAGACAGAACCATCTTACACCTT 472
   |||||
QY 143 LysGlnGluValLysPheArgIleValThrLeuPheSerAspPheLysProTyrLysThr 162
   |||||
DB 473 GGGTCCACAGTTCCTATCGATCTTCAACGTCACCAACCAAGCTGACCCGTGGGCGG 532
   |||||
QY 163 SerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGln---TyrLeuSer 181
   |||||
DB 533 ACGGTCATGTGTCAACATTGAGAACCCGGAAGGCATCCCGGTCAAGACGACTCTTGCT 592
   |||||
QY 182 GlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIleLeu 201
   |||||
DB 593 TCTAGAACCAAGCTTGCGCTTGTCCCTTGTGGGACATTCCGGAACCTGTCACATG 652
   |||||
QY 202 GlyAspTyrSerIleGlnValIleValAsnAsp-----GlnThrTyrTyrGlnSer 218
   |||||
DB 653 GGCAGATGAGAGATCCAGCTCTACTATGAAATCAACACAGCAGAGCTTCTCCACAGAG 712
   |||||
QY 219 PheGlnValSerGluTyrValLeuProLysPheGluValThrLeuGln-----ThrPro 236
   |||||

```

```

DB 713 TTGAGGTGAGAGAGTACTGCTGCCAGTTCGAGGTCACTAGTACGACCTACAGAGAA 772
QY 237 LeuTyrCysSerMetLeuSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyr 256
   |||||
DB 773 TTCTACTACATCTATTAACAGAGAGGCGCTGGAGTCAACATACCCCGAGGTTCTCTAC 832
   |||||
QY 257 GlyLysProValLysGlyAspValThrLeuThrPhe----- 268
   |||||
DB 833 GGGAGAGAAATGGAGGAACTGCTTGTATCTTGGGATCCAGAGATGGCAAGAGAG 892
   |||||
QY 269 -----LeuProLeuSerPheThrGlyLysLysAsn 279
   |||||
DB 893 ATTTCCCTGCTGAATCCCTCAAGCCCATTCGATGAGAGATGCTGGGGAGGTTGTG 952
   |||||
QY 280 IleThrLysThrPheLysIleAsnGlySerAlaAspPheSerPheAsnAspGluMet 299
   |||||
DB 953 CTGAGCCGGAAGTACTGCTGCGAGCGGGTGCACAAACCCCGACAGAAAGCTGTGGGG 1012
   |||||
QY 300 LysAsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSerProGly 319
   |||||
DB 1013 AAGTCTTGTGACGTGCTGCC----- 1033
   |||||
QY 320 ProValGluIleLeuThrThrValThrGluSerValThrGlyIleSerArgAsnValSer 339
   |||||
DB 1034 -----ACCGTCACTTGCACCTCGACGACGACATGATGTCAGGACAGCGC 1078
   |||||
QY 340 ThrAsnValPhePheLysGlnHisAspTyrIleIleGluPhePheAspTyrThrThrVal 359
   |||||
DB 1079 AGCGGATCCCACTCGTACCTCTCCCTACAGATCACTTACCAAGACACCAAGTAC 1138
   |||||
QY 360 LeuLysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsn--- 378
   |||||
DB 1139 TTCAAACACAGATGCCCTTGTGACCTCATGTGTGTGTGACGAACCTGATGCTCTCCA 1198
   |||||
QY 379 -----GlnLeuThrLeuGluIleArgArgAsnAsnValIleThrValThrGlnArg 396
   |||||
DB 1199 GCCTACCAAGTCCCGCGGACGTCAGGGCGAGACACTGTGCAAGTCTTCAACAGGGA 1258
   |||||
QY 397 AsnTyrThrGluTyrTrpSerGlySerAsnSerGlyAsnGlnLysMetGluAlaValGln 416
   |||||
DB 1259 GATGGGTGGCCAAACTCAGATCAACACACACCCCAAGCCAGCAAG--- 1303
   |||||
QY 417 LysIleAsnTyrThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGlu 436
   |||||
DB 1304 -----CCCTTGACATCAACGTCGCGACGAAGAACGACGAGCTCTCG 1345
   |||||
QY 437 AspSerSerGluLeuGlnLeuLysAlaTyrPheLeuGlySerLysSerMetAlaVal 456
   |||||
DB 1346 GAGGACAGACAG-----GTTACCGAGACCATGACGAGCTCTG 1381
   |||||
QY 457 HisSerLeuPheLysSerProSerLysThrTyrIleGlnLeuLysThr----- 472
   |||||
DB 1382 CCTACAGCACCGGTGGCAACTCAACATTTACTGTGATCTCTCACTGCTACTACAGAG 1441
   |||||
QY 473 -----ArgAspGluAsnIleLysValIleGlySerProPheGluLeuValIleSerGlyAsn 490
   |||||
DB 1442 CTCAGACCGGGGAGACCTCAACGTCACATCTCTCTGGAATGAGACCGGCCACAGAG 1501
   |||||
QY 491 LysArgLeuLysGluLeuSerTyrMetValIleSerArgGlyGlnLeuValAlaValGly 510
   |||||
DB 1502 GCCAAGATCCGTACTACATCACTGATCATGAAACAGGCGAGCTGTGGAAGCGGGA 1561
   |||||
QY 511 Lys-----GlnAsnSerThrMetPheSerLeuThrProGluAsnSer 524
   |||||
DB 1562 CGCCAGGTGAGAGCCCGGACGAGACCTGTGTGTCTGCTGCTGTCATCACACCGAG 1621
   |||||
QY 525 TrpThrProLysAlaCysValIleValIleValTyrTyr-----IleGluAspArgGly 540
   |||||
DB 1622 TTATCCCTTCTCTCGCGCTGTGGGTACTACAGCTGATGTGGTGCAGCGCCAGAG 1681
   |||||
QY 541 GluIleIleSerAspValLeuLysIleProValGln-----LeuVal 554
   |||||
DB 1682 GAGTGTGTGCGCATCTCCGTGTGTGTGACGTCAAGAGACTCTGCTGTGGCTGTGCTGTG 1741
   |||||

```

QY 555 PheLeuAsnLysLeuLeuTyrTrpSerLysValIleAlaGlu----- 569  
 Db 1742 GTAAAAAGCGGC-----CAGTCAGAAAGCCGCGACCTGTA 1777  
 QY 570 ProSerGluLysValSerLeuArgIleSerValThrGlnProAspSerIleValGlyIle 589  
 Db 1778 CCGCGGACAGATGACCTCGAAGATGAGGTGACACCGGGCCCGGGTGGTA---CTG 1834  
 QY 590 ValAlaValAspLysSerValAsnLeuMetAsnLysSerAsnAspIleThrMetGluAsn 609  
 Db 1835 GTGGCCGTGACAGAGCGCTGTTCGTGCTGAATTAAGAAACAACCTGACGAGACTAAG 1894  
 QY 610 ValValHisGluLeuGluLeuTyrAsnThrGlyTyrTyrLeuGlyMetPheMetAsnSer 629  
 Db 1895 ATCTGGAGAGTGGTGAAGAAAGCAGACATCGCTGACACCCCGGCGAGTGGAAAGATTAC 1954  
 QY 630 PheAlaValPheGlnGluCysGlyLeuTyrPvalIleuThrAspAla----- 644  
 Db 1955 GCCGGTGTCTTCTCCAGCGAGGGCTGACCTTCACGACGACAGTGGCCAGACAGCCGCC 2014  
 QY 644 ----- 644  
 Db 2015 CAGAGGGCAGAACTTAGTGCCTCCGACGACGCCGCCGACGCCGTTCCGTGACGCTC 2074  
 QY 645 -----AsnLeuThrLysAspTyrIle 651  
 Db 2075 ACGGAGAACGAAATGGAACAAGTCGGCAAGTACCCCAAGAGAGTGGCGAAGTCTGCCAG 2134  
 QY 652 AspGlyValTyrAspAsn-----AlaGluTyrAlaGluArgPheMetGlu 666  
 Db 2135 GACGGCATCGGGAGAACCCCATGAGATTCTCGTGCAGCGCGGACCGGTTTCATCTCC 2194  
 QY 667 GluAsnGlu-----GlyHisIleValAspIleHis 676  
 Db 2195 CTGGGGGAGCGCTGCAAGAGTCTTCTGACCTGCTGCACTACATCAAGAGCTGCCG 2254  
 QY 677 -----AspPheSerLeu 680  
 Db 2255 CCGGACGACCGCGCGGCGACCACTCGGCGCTGCGCAGAGATACCTGATGAGACATC 2314  
 QY 681 GlySerSerProHisVal-----ArgLysHisPheProGluThrTyrIleTrp----- 696  
 Db 2315 ATTGCAGAGAGAACATCGTTCCGAGAGTGTCCAGAGAGCTGGCGTGGAAACGTT 2374  
 QY 697 -----LeuAspThrAsnMetGlyTyrArgIleTyr 706  
 Db 2375 GAGGACTTGAAGAGCCAGCAAAAATGAAATCTCTACGAGCTC----- 2419  
 QY 707 GlnGluPheGluValThrValProAspSerIleThrSerTyrValAlaThrGlyPheVal 726  
 Db 2420 -----ATGAATATATTTTGAAGAAGCTCATACACAGCGGAGATTCTGGCTGTGACC 2473  
 QY 727 IleSerGluAspLeuGlyLeuGlyLeuThrThrThrProValGluLeuGlnAlaPheGln 746  
 Db 2474 ATGTGGAGACAAGAAAGGATC---TGTGGGACAGACCCCTTGAGGTCAAGTAAGCAG 2530  
 QY 747 ProPhePheIlePheLeuAsnLeuProTyrSerValIleArgGlyGluGluPheAlaLeu 766  
 Db 2531 GACTTCTTATCATGACCTGGCGGTACCTTCTGTGTTCGAAACAGACAGGCTGGAATC 2590  
 QY 767 GluIleThrIlePheAsnTyrLeuLysAspAlaThrGluValIleIleGlu----- 785  
 Db 2591 CGAGCCGTTCTCTACATTAAC---CGGACAGAACCAAGAGCTCAAGGTGAGGAGACTA 2647  
 QY 786 ---LysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsnAlaThrGlyHis 804  
 Db 2648 CTCACAAATCCAGCGCTTTCGAGCGCTGGCCACACACAGAGAGS-----CGTCAAC 2695  
 QY 805 GlnGlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPheProIleArgPro 824  
 Db 2696 CAGCAGACCATTAACATCCCGCCCAAGTCTCGTTGTCGTTTCATATGTATCATCGTCCG 2755

QY 825 -----ThrHisLeuGlyGluIleProIleThrValThrAlaLeuSerProThrAlaSer 842  
 Db 2756 CTAAAGACCGGCTCTCAGAGAGTGAAGTGAAGCTGCGTCTTACATCATTTCACTACGT 2815  
 QY 843 AspAlaValThrGlnMetIleLeuValLysAlaGluGlyIleGluLysSerTyrSerGln 862  
 Db 2816 GACGGTGTGAGAAAGTCTCTGAAGGTCGTGGCCGGAAGAAATCAGATGAACAAACTGCG 2875  
 QY 863 SerIleLeuLeuAspLeuThrAspAsnArgLeu-----GlnSerThrLeuLysThrLeu 880  
 Db 2876 GCTGTT---CGCACCTTGATTCAGAACGCTGGGCCCTGTAAGATGACAGAAAGAG--- 2929  
 QY 881 SerPheSerPheProProAsnThrVal-----ThrGlySerGluArg 894  
 Db 2930 -----GACATCCACCTGACAGACCTCAAGTACCAAGTCCCGGACACCGAGCTTAGACC 2983  
 QY 895 -----ValGlnIleThrAlaIleGlyAspValLeuGlyProSerIleAsnGly--- 910  
 Db 2984 AGAATTCTCTGCAAGAGAGCCCAAGTGGCCAGATGACAGAGATGCCGTGACCGCGAA 3043  
 QY 911 ---LeuAlaSerLeuIleArgMetProTyrGlyCysGlyGluGlnAsnMetIleAsnPhe 929  
 Db 3044 CCGCTGAAGACACTATGTGTGACCCCTCGGGCTCGGGGAAACAGAACTGATCGGCATG 3103  
 QY 930 AlaProAsnIleTyrIleLeuAspTyrLeuThrLysLysGlnLeuThrAspAsnLeu 949  
 Db 3104 ACGCCACAGTATGCTGCTGTGATTACTGATGAAGAAACGACATGGGAGAAAGTTCCGC 3163  
 QY 950 LysGluLys-----AlaLeuSerPheMetArgGlnGlyTyrGlnArgGluLeuLeu 966  
 Db 3164 CTAGAGAAAGCGGACAGGGGCGCTTGAGAGCTCATCAAGAAAGGGGTATCACCCAGACGCTGGCC 3223  
 QY 967 TyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyrAspProSerGlySerThr 986  
 Db 3224 TTCAGACAACCCAGCTGCTGCTTGGCGCTTGGCTTGTGAACCGGCAACC---AGCAAC 3277  
 QY 987 TrpLeuSerAlaPheValLeuArgCysPheLeuGluAlaAspProTyrIleAspIleAsp 1006  
 Db 3278 TGGCTGACCGGCTTACGTGTGTGTAAGTCTTCTCTGCGCTGTCAACCTCATCGCATCGAC 3337  
 QY 1007 GlnAsnValLeuHisArgThrTyrThrTrpLeu---LysGlyHisGlnLysSerAsnGly 1025  
 Db 3338 TCCCAAGTCTCTGCGGGGCTGTGAATGGCTGATCTCGAGAACAGACAGAACGCCACGGG 3397  
 QY 1026 GluPheTrpAspProGlyArgValIleHisSerGluLeuGlnGly----- 1041  
 Db 3398 GTCTTCCAGAGAGATGCCCGCTGATACCAAGAAATGATTTGGATTAACGAAACAC 3457  
 QY 1042 AsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArgLys 1061  
 Db 3458 AACGAGAAAGACATGCGCTTCAGCGCTTGTCTCATCTGCTG----- 3502  
 QY 1062 TyrGlnProAsnIleAspVal---GlnGluSerIleHis----- 1073  
 Db 3503 ---CAGAGAGGTAAGATATTGGAGAGAGACGTCACACAGCTCCAGGACGATCACT 3559  
 QY 1074 -----PheLeuGluSerGluPheSerArgIleIleSerAspAsnTyrThrLeu 1089  
 Db 3560 AAAGCAGAGACTTCTTGAAGCCCAATAC---ATGAACCTACAGAGATCTCTACCTGG 3616  
 QY 1090 AlaLeuIleThrTyrAlaLeuSerSerValGlySerProLysAlaLysGluAlaLeuAsn 1109  
 Db 3617 GCCATTGCTGGTATGCTTGTGCGCAATGGGACAGGCTGAAGGGGCTTTTAAACAA 3676  
 QY 1110 MetLeuThrTyrArgAlaGluGlnGluGlyGlyMetGlnPheThrPvalSerSerGluSer 1129  
 Db 3677 TTTTCGACC-----ACAGCC 3691  
 QY 1130 LysLeuSerAspSerTrpGln-----ProArgSerLeuAspIleGluValAlaAla 1146  
 Db 3692 AAAGTAAAGAACCGCTGGAGAGACCTGTGAAGCAGCTTCAACAGTGAAGGCCACATCC 3751  
 QY 1147 TyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGluGlyIle---ProIleMet 1165

```

|||||
3752 TATCCCTCTGGCC---CTACTGCAGCTAAAGACTTTGAGCTTGGCCCTCCGCTGCG 3808
Qy 1166 ArgTrpLeuSerArgGlnArgAsnSerLeuGlyPheAlaSerThrGlnAspThr 1185
Db 3809 CGTTGGCTCAATGACAGAGTACTACCGTGTGGCTGCTACCTACCGGCACTTCC 3868
Qy 1186 ValAlaLeuValAlaLeuSerGluPheAlaAla---LeuMetAsnThrGluArgThrAsn 1204
Db 3869 ATGGGTTCACAAAGCTTGCTGCATATACCAAAAGACGCCCTGACACACAGAACTGAC 3928
Qy 1205 IleGlnValThrValThrGlyProSerSer-----ProSerProLeu 1218
Db 3929 CTTGATGTGCTCCCAACTGCGACGCG--CAGCTCCAAATGATCAACCACTATCCACTG 3987
Qy 1219 AlaValValGlnProMetAlaVal-----Asn-IleLeu 1229
Db 3988 GGAATCTGCGACGCTCTCCGATCAGAAAGACCAAGAAATAGAGTTTCAACATCAC 4047
Qy 1229 ValAsnGlyPheGlyPheAlaIleCyGlnLeuAsnValValTyraAsnValValAla 1249
Db 4048 AGCTGAAGGAAAGGCGCAAGGACCTTGTCGTGGTGACATGTACATGCTAAGGCCAA 4107
Qy 1249 GglySerSerArgArgArgSerIleGlnAsnGlnValPheAspLeuAspValAl 1269
Db 4108 AGATCACTCACTGTATATAA-----TTCCGACCTCAAGGTCAC 4146
Qy 1269 aVal-----LysGluAsnLysAspAspLeuAsnHisValAspLe 1282
Db 4147 CATTAACCAACACCGGCAAGAAAGAGGCTTCAGATGCCAAGAACTATGATCTCT 4206
Qy 1282 uAsnValCyThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluValAs 1302
Db 4207 TGAATGTGTACCGAGTACCGGGAGAACAGAGATGCCATATGTCTATATGGACATATC 4266
Qy 1302 nLeuLeuSerGlyPheMetValProSerGln-----AlaIleSerLeu 1317
Db 4267 CATGATGACTGCTGTGCTTCCAGACACAGATGACCTGAGCAGCTGCGCATGTGTGA 4326
Qy 1317 rGluThrValLysLysValGluTy-----AspHisGlyLysLeuAsnLe 1332
Db 4327 CAGATACACTTCCAAAGTATGAGCTGAGACAAAGCCTTCTCCGATGAGAACCCCTCAT 4386
Qy 1332 uTyLeuAspSerValAsnGluThrGln---PheCyValAsnLysProAlaValArgAs 1351
Db 4387 CTACTGACCAAGGCTCTACACTCTGAGAGATGACTGTCTTCAAGTTCAACATA 4446
Qy 1351 nPheLysValSerAsnThrGlnAspAlaSerValSerIleValAspTyTyTyGluProAr 1371
Db 4447 CTTTAATGTAGAGCTTATCCAGCTGAGCAGTCAAGGTCTACCGCTATATCAACTGGA 4506
Qy 1371 gArgGlnAlaValArgSerTyraAsnSerGluVal-----LysLeuSerSerCyAs 1388
Db 4507 GGAAGACTGTAACCGGTTCTACCTCCGAAAGAGAGAGATGAAAGCTGAACAAG--- 4561
Qy 1388 pLeuCySerSerAspValGlnGlyCyAsArgProCyGluAsp 1401
Db 4562 -CTCTGCGCGTGTAT---GAACGTGTGCGCTGTGCTGAGAGAG 4597

```

```

RESULT 6
US-09-142-334-23
; Sequence 23, Application US/09142334
; Patent No. 6268485
; GENERAL INFORMATION:
; APPLICANT: Fatties, Timothy C.
; APPLICANT: Harrison, Richard A.
; TITLE OF INVENTION: Down-Regulation Resistant C3 Convertase
; FILE REFERENCE: 4-30443/A/IMU/PCT
; CURRENT APPLICATION NUMBER: US/09/142,334
; CURRENT FILING DATE: 1999-04-15
; EARLIER APPLICATION NUMBER: PCT/GB97/00603
; EARLIER FILING DATE: 1997-03-04
; NUMBER OF SEQ ID NOS: 35

```

```

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 23
; LENGTH: 5067
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-142-334-23

```

```

Alignment Scores:
Pred. No.: 4,46e-88 Length: 5067
Score: 931.00 Matches: 381
Percent Similarity: 40.88% Conservative: 287
Best Local Similarity: 23.32% Mismatches: 609
Query Match: 12.67% Indels: 358
DB: Gaps: 63

```

US-10-020-095-4 (1-1428) x US-09-142-334-23 (1-5067)

```

Qy 3 GlyProProLeuLeuThrAlaAlaHisLeuLeuCyValCyThrAlaAlaLeuAlaVal 22
Db 76 GGTCCAGCCTGCTG-----CTCCGCTACTAACCACCACTCCCGGCTCTG 123
Qy 23 AlaProGlyProArgPheLeuValThrAlaProGlyTyleIleArg----- 37
Db 124 GGG---AGTCCCATGACTTATCATCACCCCAACATCTTGCGCTGAGAGCGAGAG 180
Qy 38 -----ProGlyGlyAsnValThrIleGlyValGluLeu 49
Db 181 ACCATGATGTGTGAGAGGCCACAGCGGCAAGGAGATTCAGTCACTGTACTGTCCAC 240
Qy 50 GluHis-----CyProSerGlnValThrValValAlaGluLeu 63
Db 241 GACTTCCAGCAAGAAACTAGTGTCTGTCAGTGAAGAGCTGTGCTGACCCCTGCGACC 300
Qy 64 LysThrAlaSerAsnLeuThrValSerVal---LeuGluAlaGluValPheGluLys 82
Db 301 AACCAATGGGCAAGCTCACCTTACAGATCCAGCAACAGGAGTTCAAGTACAGAAAG 360
Qy 83 GlySerPheTyThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGluIleTy 102
Db 361 GGGCGCAACAAGTTCGTGACCGTGCAGGCC-----ACCTCGGAGCCCAAGTGTG 411
Qy 103 GlnLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnThrArgLeu 122
Db 412 GAGAGAGGTGTGTGCTGTCACCTGACAGCGGCTACCTC----- 450
Qy 123 SerPheGluThrLysArgLysSerValPheIleGlnThrAspLysAlaLeuTyTyPro 142
Db 451 -----TTCATCAAGACAGACAAAGCCATCTACACCCCT 483
Qy 143 LysGlnGluValLysPheArgLysIleValThrLeuPheSerAspPheLysProTyTyThr 162
Db 484 GGTCCACAGTTCTCATGCGATCTTCAACCGTCAACCAACAGCTGTACCGGAGCCGG 543
Qy 163 SerLeuAsnLysLeuLysAspProLysSerAsnLeuIleGlnGln---TrpLeuSer 181
Db 544 ACCGTATGTCAACATTTGAGAACCCGGAAGGAGCATCCGCTCAAGCAGACTCTTGTCT 603
Qy 182 GlnGlnSerAspLeuGlyValLysSerLysThrPheGlnLeuSerSerHisProIleLeu 201
Db 604 TCTCAAGAACAGTGTGCGCTGTGCGCTTGTCTTGGAGACATTCGGAACCTCGTCAACATG 663
Qy 202 GlyAspTrpSerIleGlnValGlnValAsnAsp-----GlnThrTyTyGlnSer 218
Db 664 GGCAGTGGAAAGATCCGAGCTTACTATGAAGTCAACACAGAGAGCTTCTCAGATGAG 723
Qy 219 PheGlnValSerGluTyValLeuProLysPheGlnValThrLeuGln-----ThrPro 236
Db 724 TTGAGGTGAAGAGATGACGTGCGCCAGTTTCAAGTCAATAGTGAAGCCTTACAGAGAA 783
Qy 237 LeuTyTyCySerMetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyTyThrTy 256
Db 784 TTCTACTAATCATATTAAGAGAAAGGCGCTGAGAGTCAACATCACCCGAGTTCTCTAC 843

```

QY	257	GIlyrsProValLyGIyAspValThrLeuThrPhe-----	268
Db	844	GGGAAGAAAGTGGAGGAACTGCTTTTGTCATCTTCGGAGTCCAGAGTGCAGAACAGAGG	903
QY	269	-----LeuProLeuSerPheTyrPoliYslylsYlsaaN	279
Db	904	ATTTCCTGCTCGAATCCCTCAAGGCGCATTCGGATTGAGAGATGGCTCGGGAGAGTTTGG	963
QY	280	ILethrlysrthPheLysILEaNGlySerIlaaSPheSerPheaaNspgluIuMet	299
Db	964	CTGAGCCGGAAAGTACTGCTGGACGGGGTGGAGAACCCTCCGAGCAGAAAGACCTGGTGGGG	1022
QY	300	LysaaNValMetaspSerSeraaNglyLeuSerGIuTyrrLeuaspLeuSerSerProgly	319
Db	1024	AAGTCTTTGTAGCTGTCTGC-----	1044
QY	320	ProValGIuILEuThrValThrgIuSerValThrgIyILEseraNgaaNValSer	339
Db	1045	-----ACCGTACTCTTGGACTCACTGAGGAGGACGATGGTGACGAGCAGAGCC	1081
QY	340	ThraaenValPhePheLysGIaNHIsaPyrILEleGIuPhePheaaPyrThrThrVal	359
Db	1090	AGCGGATCCCATGTCACCTCCCTCAAGATCACTTCCACCAAGACCCCAAGTAC	1144
QY	360	LeuLyPProSerLeuaspNheThralaThrValLysValThraGlaaapglYaen---	378
Db	1150	TTCAAAACCAAGAAATCCCTTGACCTCATGGTGTTCGAGCAACCCTGATGGCTCTCA	1201
QY	379	-----GIuLeuThrLeuGIuIuNgaaNaaenValIILEthryalThrgIaaNg	396
Db	1210	GCCTAACCAAGTCCCGTGGCACTCCAGGGCGAGAACCTGTCAAGTCTTAAACCCAGGGA	1261
QY	397	AsnTyrrthGIuTyrrTyrPserLySeraaNserGIyaaNGIuIuMetGIuIaValaGIa	416
Db	1270	GATGGCGTGGCCAAATCACTCAACAACAACCCAGCAGAAAG-----	1311
QY	417	LysILEaSnTyrrThValProGIuSerGIyThrPheLysILEuIuPheProILEuGIu	436
Db	1315	-----CCCTTGAGCATCAACGGTGGCGCAGAAAGACAGAGACTCTCG	1351
QY	437	AspSerSerGIuLeuGIuIuLeuLysalATyrrPheLeuGIySerLySerSerMetIaVal	456
Db	1357	GAGGCAAGACAGCA-----GCTAACAGGACCATGACAGGCTCTG	1391
QY	457	HisSerLeuPheLysSerProSerLySrThyrrILEGIuLeuLySrThr-----	472
Db	1393	CCCTACAGACCCGTGGGCACTTCCAACAATTACCTGCATCTTCAAGTCTTACTGACAGG	1451
QY	473	-----ArgaspGIuaspNILEySValGIySerProPheGIuLeuValaValSerGIyaaN	490
Db	1453	CTCAGACCCGGGAGACCTCAACGTCATCTTCCTCGCAATGAGACGGCGCCACGAG	1511
QY	491	LysaIaGIuLeuLySGIuLeuSerTyrrMetValaValSerArgILEuIuValaIaValGIy	510
Db	1513	GCCAAAGATCCGCTACTACACTTACATCAATGATCAAGAACAGGGGAGGCTGTGGAAGGGGGA	1571
QY	511	Lys-----GIuaspSerThrMetPheSerLeuThrProGIuaspSer	524
Db	1573	CGCCAGGTGCGAGAGCCGGCCAGAGACTGTGGTGGCTGCCCTGTGCATCAACCCAGAC	1631
QY	525	TrpThrProLySaIaCySValILEValTyrr-----ILEGIuaspApGIy	540
Db	1633	TTCAATCCCTTCTTCCGCTGGTGGGTACTCAACAGCTATCGATCGATCGATCGAGGCGCAGAG	1691
QY	541	GIuILEILESeraspValLeuLySILEProValGIa-----LeuVal	554
Db	1693	GAGGTGGTGGCCGACTCCGTGGGTGGGTGACGTCAAGACACTCGTGGCTGGCTGGTGGTGG	1751
QY	555	PheLySaSnTyrrILEySLeuTyrrTyrSerLySValLySaIaGIu-----	569
Db	1753	GTAAGAAAGGCGC-----CACTCAAGAAACCGGACCGCTGTA	1781
QY	570	ProSerGIuLySValSerLeuArgILESerValThrgIuProaspSerILEValaGIyLE	589

[illegible]

```

Db      2827 GACGGTCAGAGACTCCCTGAAGTCCGTCCCGGAGAGATCAGATGACAAAACCTGTG 2886
Qy      SerileuLeuAspLeuThrAspAsnArgLeu-----GlnSerThrLeuLeu 880
Db      2887 GCTGTT---CGACCCCTGGATCCAGAACGCTGGGCCCTGAAGAGATGACAAAAGAG--- 2940
Qy      SerPheSerPheProProAsnThrVal-----ThrGlySerGluArg 894
Db      2941 -----GACATCCCACTGCAGACCTCAGTACCAAGTCCCGGACACCGAGTCTGAGACC 2994
Qy      895 -----ValGlnIleThrAlaIleGlyAspValIleGlyProSerIleAsnGly--- 910
Db      2995 AGAATTCCTCTGCAGAGGAGCCCAAGTGGCCCAAGATGACAGAGATGCCGTGACGCGAA 3054
Qy      911 ---LeuAlaSerLeuIleArgMetProTyrGlyCyseGlyGluGlnAsnMetIleAsn 929
Db      3055 CGCGTGAAGCACCTCATTTGTGACCCCTCGGGCTCGGGGAAACAGAACATATATCGGCATG 3114
Qy      930 AlaProAsnIleTyrIleLeuAspTyrLeuThrIlyblyblyblyGlnLeuThrAspAsn 949
Db      3115 ACGCCACGATCATCGTGCATTAACCTGGATGAAACGAGACAGTGGAGAAAGTCCGCC 3174
Qy      950 LysGlyblyblyblyblyblyblyblyblyblyblyblyblyblyblyblyblybly 966
Db      3175 CTAGAGAAAGCGGACGAGGCGCTTGAGCTCATCAAGAAAGGAGGTCACCCACACCTG 3234
Qy      967 TyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyrAspProSerGlySerThr 986
Db      3235 TTCAGACAAACCCAGCTGCTGCTTTCGGCTTCGCGAACCGGACCC-----AGCAC 3288
Qy      987 TrpLeuSerAlaPheValLeuArgCysePheLeuGluAlaAspProTyrIleAspIleAsp 1006
Db      3289 TGGCTGACCGGCTCAAGGTCAGAGTCTTCTCTGCGCTCAACCTCATCGCATCGAC 3348
Qy      1007 GlnAsnValIleuHlaArgThrTyrThrProLeu---LysGlyHlaGlnLysSerAsnGly 1025
Db      3349 TCCCAAGTCTCTCGGGGCTGTTAAATGCGTGAATCTCGAGAAAGCAAAACCCGACGG 3408
Qy      1026 GluPheTrpAspProGlyArgValIleHisSerGluLeuGlnGly----- 1041
Db      3409 GTCTTCCAGAGAGATGCGCCCTGATACACCAAGAAAGATTTGGATTCGGAACAC 3468
Qy      1042 AsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArgLys 1061
Db      3469 AAGCAGAAAGACATGCGCTCAGCGCTTGTCTCATCTCGCTG----- 3513
Qy      1062 TyrGlnProAsnIleAspVal---GlnLysSerIleHis----- 1073
Db      3514 ---CAGAGAGCTAAAGATATTGCGAGAGAGAGCTCAACAGCTGCCAGGACATCACT 3570
Qy      1074 -----PheLeuGlnSerGluPheSerArgLysIleSerAspAsnTyrThrLeu 1089
Db      3571 AAAGCAGAGACCTTCTTAAGCCCACTAC---ATGAACCTACAGAGATCTTACACTGTG 3627
Qy      1090 AlaLeuIleThrTyrAlaLeuSerSerValGlySerProLysAlaLysGluAlaLeuAsn 1109
Db      3628 GCAATGCTGCTAGTCTGCTGCGCCAGATGGCAGAGCGGAGGCGCTTCTTAACAAA 3687
Qy      1110 MetLeuThrTrpArgAlaGluGlnGlyLysMetGlnPheTrpValSerSerGluSer 1129
Db      3688 TTTCTGACC-----ACAGCC 3702
Qy      1130 LysLeuSerAspSerTrpGln-----ProArgSerLeuAspIleGluValAlaAla 1146
Db      3703 AAAGATAGAAACCGTGGAGAGACCTGTGAAGCCTCTCAACAGTGGAGCCACATCC 3762
Qy      1147 TyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGluGlyLe---ProIleMet 1165
Db      3763 TATGCCCTCTTGCC---CTACTGACGCTAAAGACCTTGACTTGTGCGCTCCCGTGGTG 3819
Qy      1166 ArgTrpLeuSerArgGlnArgAsnSerLeuGlyGlyPheAlaSerThrGlnAspThrThr 1185
Db      3820 CGTTGGCTCAATGAACAGAGATACTACGCTGTGCTATGCTCTACCCAGGACCACTTC 3879

```

```

Qy      1186 ValAlaLeuLysAlaLeuSerGluPheAlaAla---LeuMetAsnThrGluArgThrAsn 1204
Db      3880 ATGATGTTCAAGCCTTGCTCAATACCAAAAGACGCCCTTGACCAACGAGAACTGAAAC 3939
Qy      1205 IleGlnValThrValThrGlyProSerSer-----ProSerProLeu 1218
Db      3940 CTGATGATGCTCCCAACCTGACCGCG-CAAGCTCCAAAGATCAACCAACGATTCACATG 3998
Qy      1219 AlaValValGlnProMetAlaVal-----Asn-IleSer 1229
Db      3999 GGAATTCGACGCTCTCTGATCAGAGACCAAGAAATGAGCGTTCCACATGAC 4058
Qy      1229 AlaAsnGlyPheGlyPheAlaIleCyseGlnLeuAsnValTyrAsnValLysAla 1249
Db      4059 AGCTGAAGAAAGGCGCAAGGACGACCTTGCGGTGGTGAACAATGATCATGCTGAAGGCA 4118
Qy      1249 rGlySerSerArgArgArgSerIleGlnAsnGlnGluAlaPheAspLeuAspValAl 1269
Db      4119 AGATCAACTCACCTGTAATMA-----TTGACCTCAAGGTAC 4157
Qy      1269 aVal-----LysGluAsnLysAspAspLeuAsnHisValAspLe 1282
Db      4158 CATAAACACGACCGGAAACAGAAAGAGGCTCAGATGCCAAGACATATGATCT 4217
Qy      1282 uAsnValCyseThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluValAs 1302
Db      4218 TGAGATCTGACAGGATCCGCGGAGACCAAGATGACATATGATATGACATATG 4277
Qy      1302 nLeuLeuSerGlyPheMetValProSerGlu-----AlaIleSerLeu 1317
Db      4278 CAGATGACACGCTGCTTGTCTCCAGACACAGATGACCTGAAGCAGCTGGCCCAATGCTTGA 4337
Qy      1317 rGluThrValLysValGluTyr-----AspHisGlyLysLeuAsn 1332
Db      4338 CAGATACATCTCAAGATATAGCTGACAAAGCTTCTCCGATAGGAACACCTCATCAT 4397
Qy      1332 uTyrLeuAspSerValAsnGluThrGln---PheCyseValAsnIleProAlaValArgAs 1351
Db      4398 CTACCTGAGACAGAGTCTACACCTGAGGATGATCTGTACGTTCAAGTTCAACAAATA 4457
Qy      1351 nPheLysValSerAsnThrGlnAspAlaSerValSerIleValAspTyrTyrGluProAr 1371
Db      4458 CTTTAATGTAGACTTATCCAGCTGAGCAGCAAGTCTACGCTATTACAACTGGA 4517
Qy      1371 gArgGlnAlaValArgSerTyrTrpAsnSerGluVal-----LysLeuSerSerCyse 1388
Db      4518 GGAAGCTGACCGGTTTCAATCCGGAAGAAAGAGATGGAAGCTGAACAG----- 4572
Qy      1388 PLeuCyseSerAspValGlnGlyCyseArgProCyseGluAsp 1401
Db      4573 -CTTGCCTGAT---GAAGTGTGCCGCTGTCTGAGAGAG 4608

```

```

RESULT 7
US-08-662-227-1
Sequence 1, Application US/0866227
Patent No. 5922320
GENERAL INFORMATION:
APPLICANT: VOGEL, CARL-WILHELM
APPLICANT: BREDEHORST, REINHORST
APPLICANT: KOCK, MICHAEL
APPLICANT: FRITZINGER, DAVID
TITLE OF INVENTION: RECOMBINANT PROCVF
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLOH, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:

```



```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/662,227
FILING DATE: 14-JUN-1996
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 1126-0107-0X
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEO ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5948 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-662-227-1

Alignment Scores:
Pred. No.: 5.24e-84 Length: 5948
Score: 894.00 Matches: 361
Percent Similarity: 40.90% Conservative: 295
Best Local Similarity: 22.51% Mismatches: 672
Query Match: 12.17% Indels: 276
DB: 2 Gaps: 56

US-10-020-095-4 (1-1428) x US-08-662-227-1 (1-5948)
QY 18 AlaAlaLeuAlaVal-----AlaProGlyProArgPheLeuValThrAla 32
DB 31 GCGCTCTATGATGATGTTTCCAGGCTCTTCATGAGGCTCTCAACCTCATCACC 90
QY 33 ProGlyIleIleArgProGlyGlyAsnValThrIleGlyValGluLeuLeu---GluHis 51
DB 91 CCTGCTGTTTGGCAACAGACAGACAGAAAGCAATTTTGGTGGAGGCCCATGAGCACT 150
QY 52 CysProSerGlnValThrValIleValysala-----GluLeuLeuLysThrAlaSer 67
DB 151 ACTCCAAACAGCTGATGATGATCCAGCAGAGGAGGATCTTGTCAC-----CCAAC 210
QY 68 AsnLeuThrValSerValLeuGluIleGluGlyValPheGluLysSerPheLysThr 87
DB 211 CAACCAAGATGATGATGATGATCCAGCAGAGGAGGATCTTGTCAC-----CCAAC 261
QY 88 LeuThrLeuProSerLeuProLeuAsnSerAlaAspGluIleTyrGluLeuArgValThr 107
DB 262 ATTAGAGATTCACAGAAAGAGAGTACGAGCTCCAGGCAAAATCAATATGATGTTG 321
QY 108 GlyArgThrGlnAspGluIleLeuPheSerAsnSerThrArgLeuSerGlnIleSerPheGlu 127
DB 322 CAAGTAACTGCTCCTCAAGTGAATGAAAAGTGTCTCTCTTTCAACAGAGTAC 381
QY 128 ArgIleSerValPheIleGlnThrAspIleValLeuTyrLysProLysGlnIleValLys 147
DB 382 -----TTTCTGTTTCCAGACAGATAAAGCATCTATACACAGAGGCTCCAGATCTC 435
QY 148 PheArgIleValThrIleuPheSerAspPheLysProTyrLysThrSerLeuAsnIleLeu 167
DB 436 TATCGGTTTTCATGATGATCAACACAGACAGATGATCAACAACTGTGATGTTGAG 495
QY 168 IleLysAspProLysSerAsnLeuIleGlnIleTyrLeuSerGlnIleSerAspLeuGly 187
DB 496 TTTCAACATCCAGAGAGCATCTCTTGTCACT-----TCTAATTCATGATGACCTAAC 546
QY 188 ValIleSerLysThrPheGluLeuSerSerHisProIleLeuGlyAspTyrSerIle--- 206
DB 547 TTCTTC---TGCGCTTACAAATTTACAGACCTTGTGAGTTGGGAGCTTGAGAGATTGTG 603

```

```

QY 207 ---GlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGlnValSerGluTyrVal 225
DB 604 GCCAAATATGAACATCTCCAGAGAAATATACGTATATTTGATGATCAGAAATATGTG 663
QY 226 LeuProLysPheGluValThrLeuGlnThrPro-----LeuTyrCysSerMetAsn 242
DB 664 TTGCCAAGCTTGAAGTCCGCTGCAACATCAGAAAGTTTATCATTTGACCGCAAT 723
QY 243 SerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProValIleGly 262
DB 724 ---GAAATTTCCAGTGTCTATCACTCAGAGTACTTGATGAGAGAAAGTGAAGGT 780
QY 263 AspValThrLeuThrPheLeuProLeuSerPheTyrGlyLysLysAsnIle----- 280
DB 781 ---GTGGCTTGTTCCTCTTGGAGTGAATAATAGATGATGCTTAAAAAGATTTCCAGAC 837
QY 281 -----ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGlu 298
DB 838 TCACCTCAGAGAAATCCGATTTATGATGAGAGAAAGCAACACATAAAGAGATCA 897
QY 299 MetLysAsnValMetAspSerSerAsnGlyLeuSerGluTyrLysAspLeuSerPro 318
DB 898 TTCCGTTCTCATTTCCAAATCTCAATGAGCTTGTGGCAATCTCTGTATGACATCT--- 954
QY 319 GlyProValGluIleLeuThrThrValThrGlnSerValThrGlyIleSerArgAsnVal 338
DB 955 -----GTAAACGTATGACAGATCAAGATCAAGCAGATGATATGTAGTACCTGAG 999
QY 339 SerThrAsnValPhePheLysGlnHisAspTyrIleIleGluPheAspTyrThrThr 358
DB 1000 CAAGCGGATTCATATTTGTGGCATCTCCATATGATCCATCCATCAAAACCCCAAA 1059
QY 359 ValLeuLysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsn 378
DB 1060 TATTTCAAGCCAGAGATCCATATGATGAGCGGTATGTTACCAACCTGATGCTCA 1119
QY 379 GlnLeuThrLeuGluGluArgAsnAsnValValIleThrValThrGlnArgAsnTyr 398
DB 1120 CCAGCT-----GCCATGTGCCAGTGTGATACAGAGCCCTTT 1155
QY 399 ThrGluTyrTyrSerGlySerAsnSerGlyAsnGlnLysMetGluAlaValGlnLysIle 418
DB 1156 -----CATTCATGGGAACCATTTGAGATGAGATGGAGCTGTAAGCTC 1197
QY 419 AsnTyrThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGluAspSer 438
DB 1198 ATCTGAACATACCATGTAATGCTCA-----ACCTTACCAATCATCTGTTAAGACT 1248
QY 439 SerGluLeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSer 458
DB 1249 AACCATGAGAACCTCCAGAGAAACCCAGCAACAAAGTCAATGACAGCATGACCTAC 1308
QY 459 LeuPheLysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLys 478
DB 1309 CAACCCAGGAGAGATTCGAAACTATCTTATGATGAGCATTAATCACTACAGATTAAG 1368
QY 479 ValGlySer-----ProPheGluLeuValSerGlyAsn-----LysArg 492
DB 1369 CCCGAGATTAATCTTACCTGCAATTTCAATGTGAAGGCAATGCAATTCACCTAGAGCAG 1428
QY 493 LeuLysGluLeuSerTyrMetValValSerArgGlyGlnLeuValAlaValGlyLys--- 511
DB 1429 ATCAAAATTTCCACTCATATGATGATAAAGGAAAGATTTTCAAGTGTGGAGGCA 1488
QY 512 -----GlnAsnSerThrMetPheSerLeuThrProGluAsnSerTyrThr 526
DB 1489 CCCAGAGAGATGGGCAAAATCTGTGATCAATGATGATGATCAATTCACCTCAGATCTATC 1548
QY 527 ProLysAlaCysValIleValTyrTyrIleGluAspAspGlyGluIleIleSerAspVal 546
DB 1549 CTTTCTTCGCGTTGTGCTTACTACCAAGTGGGAAACAGAAATGTTGGCTGATCTCT 1608

```

QY 547 LeuYsIleProValGln-----LeuValPheIys-----AsnIys 558  
 Db 1609 GTCGGGCGTGCATGTAAGCAATACCTGCATGGAAAGTGTGTAAGAAAGACAAATCTA 1668  
 QY 559 IleYsIleuYrTrpSerIysValIysIleGluProSerGluIysValSerLeuArgIle 578  
 Db 1669 ATCAAAATG-----CCAGAGCTGCATATAAATCAAAATG 1704  
 QY 579 SerValThrGlnProAspSerIleValGlyIleValAlaValAspIysSerValAsnIeu 598  
 Db 1705 GAA---GGGGAATCCAGGTCGTGGTGTGCTGTGGCTGTGGCAAGACGATATGTT 1761  
 QY 599 MetAsnIaSerAsnAspIleThrMetGluAsnValAlaHsGluLeuGluLeuTrsn 618  
 Db 1762 CTCATGATATAATATAATAGATTAGCCAAAGATATGGGACACAAATAGAAAGATGAC 1821  
 QY 619 ThrGlyYrTrpLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCyGlyIeu 638  
 Db 1822 TTGGGCTGTACAGCTGGCAGTGGCCAGATATATCTGGGTGTGGTGAAGATGTGACTG 1881  
 QY 639 TrpValIeuThrAspAlaAsnIeu---ThrIysAsp----- 649  
 Db 1882 GCTGTGACACCCAGACTAATCTCAACACCAAGAGATCAGCTGCAAGTGTCTGAC 1941  
 QY 649 ----- 649  
 Db 1942 COTGCAATCGAGGCGTGCAGATTCTGTGTTGCTGTTGACAGACAGCAAGAAAGCG 2001  
 QY 649 ----- 649  
 Db 2002 GCAGAAATTCAGGATCAAGACCTCGTAATGCTGTGAAGATGCATGATGAGAACCC 2061  
 QY 649 ----- 649  
 Db 2062 ATGGGGTACCTTGTGAAAAGCGTGCAAAATATCATCCAGAGGAGATGCTTGAAGCT 2121  
 QY 650 -----TyrIleAspGlyValIyAsp---AsnAlaGluTyrAla 661  
 Db 2122 GCCTTCCTTGAATGCTGTGCTACATCAAGGGGGTCCAGATGAAAACCAACGGGAGAGC 2181  
 QY 662 GluArgPheMet-----GluGluAsnGlu---GlyHisIleValAspIleHisAspPhe 678  
 Db 2182 GAGTGTTCCTGGCAAGAGATGATATAGAAAGTGTTCATGACAGATGATGATATCATC 2241  
 QY 679 SerLeuGlySerSerProHisValArgYsHisPheProGluTrpIleTrpLeu--- 697  
 Db 2242 TCA-----AGGTCTGATTTCCCAAGAGTGTGTGCTGACA 2280  
 QY 698 -----AspThrAsnMetGlyTyrArgIleTyrGlnGluPheGluValThr 712  
 Db 2281 AAGGACTTGACCGAGAGCTTACGTCAGAGGAAATTCAGACAGACAAATGCTTTTAT 2340  
 QY 713 ValProAspSerIleThrSerTrpValAlaThrGlyPheValIleSerGluAspLeuGly 732  
 Db 2341 CTGAGGAGTTCATCACAACTGGGTGTGCTGCTGATACCTTACACCCACCAAGGG 2400  
 QY 733 LeuGlyLeuThrThrProValGluLeuGlnAlaPheGlnProPhePheIlePheIeu 752  
 Db 2401 ATC---TGTGTGCTGAACTTATGAAATTAAGAGTCATGAAGCTTTTATATGATCTT 2457  
 QY 753 AsnIeuProYrSerValIleArgGlyGluGluPheAlaLeuGluIleThrIlePheAsn 772  
 Db 2458 CAATGCCATATTCAGTAGTAAGATAGACGAGATTCAGATTCATTCGACACAC 2517  
 QY 773 TyrLeuYsAspAlaThrGluValIysValIleIleGluYsSerAspIysPheAspIle 792  
 Db 2518 TACGTTAAGCAGAGATATTATGTGCGAGTGAAGCTGATATCAACCCAGCCTTC----- 2571  
 QY 793 LeuMetThrSerSerGluIleAsnAlaThrGlyHisGlnGlnThrLeuLeuValProSer 812  
 Db 2572 -----TGCAGTGTCTTCCAAAAGAGACAAAGATACCCAGACAGTTCCTCAATTAAGCC 2625  
 QY 813 GluAspGlyAlaThrValIeuPheProIleArgProThrHisLeuGly-----GluIle 830

Db 2626 CTGTCTCCAGACAGTACCGGTTGTGATGTCCATTAGACGAAGATTCATGATGTT 2685  
 QY 831 ProIleThrValThrAlaIeuSerProThrAlaSerAspAlaValThrGlnMetIleIeu 850  
 Db 2686 GAGATTAAACCAAGTGTCCAGGAAGCCTTGTGTGACGCGTGTGAGGAAGAACTGAA 2745  
 QY 851 ValIysIleGluGlyIleGluYsSerYrSerGlnSerIleLeuLeuAspLeuThrAsp 870  
 Db 2746 GTTGACTCAAGGGGTAGGAATCATGTGATGATTTGTTAACTGACCAAGGCA 2805  
 QY 871 AsnArgLeuGlnSerThr-----LeuYsThrIeuSerPheSerPhe 884  
 Db 2806 AAGAGCTGTGTGACACAGCTAGAGTGAAGTCAAAAGCCGAAATTAAGTGAAGAGTG 2865  
 QY 885 ProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIleGlyAspValIeu 904  
 Db 2866 CTGTACACAGAAATTAAGAACCAAGATTTATCATCCAGGTGACCTGTGCTGATATT 2925  
 QY 905 GlyProSerIleAsnGly-----LeuAlaSerIleuIleArgMetProYrGlyCyGly 922  
 Db 2926 GAAACTCAATTTAGTGAAGTAACTCAACCATCTATTTCACTGCTGTGGCTGTGG 2985  
 QY 923 GluGlnAsnMetIleAsnPheAlaProAsnIle-----TyrIleLeuAspYrLeu 939  
 Db 2986 GAGCAAAATATGATCCGATGCGCCAGCCAGTTATTCACCTATCACTGACACCCACA 3045  
 QY 940 ThrIleYsIysGluIleuThrAspAsnIleuYsGluYsAlaIeuSerPheMetArgGln 959  
 Db 3046 GAGCAGTGGAGACTCTCGGATTAATCGCAGACTGMACTGTCAATCATGATCGTACT 3105  
 QY 960 GlyTyrGlnArgGluLeuLeuYrGlnArgGluAspGlySerPheSerAlaPheGlyAsn 979  
 Db 3106 GGTATGCCAGAGATGCTGTACAAAGAAAGCATATTCATGAGCATTTTACAAAC 3165  
 QY 980 TyrAspProSerGlySerThrTrpLeuSerAlaPheValIeuArgCyPhe---LeuGlu 998  
 Db 3166 -----CGTCACTGTGTTCTTGTGCTTACAGCATATGCTTAAAGCTTTGCCATGCT 3219  
 QY 999 AlaAspProYrIleAspIleAsnValIeuHisValArgThrYrThrTrpLeu--- 1017  
 Db 3220 GCCAAAATGCTAGCAGCATTAAGTATGAATATTTGTGAGGTGAGGTGAGTGTAT 3279  
 QY 1018 LysGlyHisGlnYsSerAsnGlyGluPheTrpAspProGlyArgValIleHisSerGlu 1037  
 Db 3280 CTGACACGCAACACCAAGATGAGGTTCAAGAAATGCCCTGTACTTTGGAACA 3339  
 QY 1038 LeuGlnGly-----GlyAsnYsSerProValThrIeuThrAlaYrIleValThr 1054  
 Db 3340 ATGACAGGAGAAATTCAGAGTGTGTAAGAAAGATATTTAACAGCTTTCATTCGTT 3399  
 QY 1055 SerLeuLeuGlyYrArg-----LysTyrGlnProAsnIleAsp-----Val 1068  
 Db 3400 GCGTGTGTAATCCAAACAAATCTGCATGACTATGTCAATGTCTAGACAGACATC 3459  
 QY 1069 GlnGluSerIleHisPheLeuGluSerGluPheSerArgGlyIleSerAspAsnYrThr 1088  
 Db 3460 AAGAGGCCCAAAATTTATCTCAAAAAGTATGAGAAA---CTGCAAGGCCCTTACAT 3516  
 QY 1089 LeuAlaLeuIleThrTyrAlaIeuSerSerValGlySerProYsAlaYsGluAlaLeu 1108  
 Db 3517 ACAGCCCTCACACCTATGCTTGGCT-----CCTGACAGACCAATC 3558  
 QY 1109 AsnMetLeuThrTrpArgIleGluGlnGluGlyMetGlnPheThrValSerSerGlu 1128  
 Db 3559 AATGATGACAGGTAATCATGAGCATCAACAGGAAGGATCATTTG----- 3606  
 QY 1129 SerIysLeuSerAspSerTrpGlnProArgSerLeuAspIleGluValAlaIaIaYrAla 1148  
 Db 3607 -----GAAATATCAATGCTCACACCCACCAACATTTGAAGGCACTTCTATGCC 3654  
 QY 1149 LeuIeuSerHisPheLeuGlnPheGlnThrSerGluGlyIleProIleMetArgTrpLeu 1168



Qy	128	ArgLiesValPheIleGlnThrAspValAlaLeuTyrLysProLysGlnGluValys	147
Db	382	-----TTCTGTTTATTCAGACGATTAAGGACCTTAACACCGGGTCTCCAGTACTC	435
Qy	148	PheArgIleValThrLeuPheSerAspPheLysProTyrLysSerLeuAsnIleLeu	167
Db	436	TATCGTGTGTTTTCTATGATGCACAAACAGCAAGATGAAACAAATGCTGATTGTTGAG	495
Qy	168	IleLysAspProLysSerAsnLeuIleGlnGlnTrpLeuSerGlnGlnSerAspLeu	187
Db	496	TTTCAGACTCCAGAAAGGCATTTCTGTCACT-----TCTAATTCACTTGACCTTAAC	546
Qy	188	ValIleSerLysThrPheGlnLeuSerSerHisProIleLeuGlyAspTrpSerIle--	206
Db	547	TTCTTCC-----TGGCCTTACAAATTTCACAGACCTGTGCAGTTGGGGACCTTGAGGATTGTG	603
Qy	207	--GlnValGlnValAlaAsnAspGlnThrTyrTyrGlnSerPheGlnValSerGluTyrVal	225
Db	604	GCCAAATATGACATTTCCCCAGAGATTATCTGCATATTTTGATGTGAGAAATATGTG	663
Qy	226	LeuProLysPheGlnValThrLeuGlnThrPro-----LeuTyrCysSerMetAsn	242
Db	664	TTGCCACACTTTGAAGTCCGTCTGCACACCATCAGAGAAGTTTTTTATCATTTGACGGCAT	723
Qy	243	SerLysHisIleAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProValLysGly	262
Db	724	---GAAATTTTCCACGTCCTATCACTCCAGAGTCACTGTGTATGGAGAGAAAGTGGAAAGT	780
Qy	263	AspValThrLeuThrPheLeuProLeuSerPheTrpGlyLysIleValAsnIle-----	280
Db	781	---GTGGCCTTGTCTCTTTGGAGTGAATAATAGATGATGCTATTAATAAAGATTTTCCAGAC	837
Qy	281	-----ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGlnGlu	298
Db	838	TCACTCACAGAGAAATTCGATTAATTGATGAGATGGGAAAGCAACATAAAAGAGATACA	897
Qy	299	MethylAsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSerSerPro	318
Db	898	TTTCGTTCTCGATTTCCAAATCTCAATAGACTGTGTGGCACTACTCTGTAGCACT---	954
Qy	319	GlyProValGluIleLeuThrThrValThrGlnSerValThrGlyIleSerArgAsnVal	338
Db	955	-----GTACACATCATGACACGATGACGACGATGATGTACTGACTGAC	999
Qy	339	SerThrAsnValPhePheLysGlnHisAspTyrIleIleGluPheAspTyrThrThr	358
Db	1000	CAAAAGCGGCATCATATTGTGGCATCTCCCTATCGATCCACTTCACAAAAAACCCTCCAAA	1058
Qy	359	ValLeuLysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsn	378
Db	1060	TATTTCAAGCCGAGATGCAATATACTGACGGGTGTATGTATCCAAACCTGATGGCTCA	1119
Qy	379	GlnLeuThrLeuGlnGluIleArgArgAsnAsnValIleThrValThrGlnArgAsnTyr	398
Db	1120	CCAGCT-----GCCCATGTGCCAGGTGTATCAGAGCCTTT	1155
Qy	399	ThrGluTyrTrpSerGlySerAsnSerGlyAsnGlnLysMetGluAlaValGlnLysIle	418
Db	1156	-----CATTCATGGGAAACCACTTTGAGTGAAGGACCTGTAAAGCTC	1197
Qy	419	AsnTyrThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGluAspSer	438
Db	1198	ATCCCGAAACATACCATTTGAATCTCAA-----AGCCTACCAATCACTGTTGAACCT	1248
Qy	439	SerGluLeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSer	458
Db	1249	AACCATGAGACCTCCCAAGAGAAAGCCAGGACCAAAAGTCCATGACAGCCATAGCTTAC	1308
Qy	459	LeuPheLysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLys	478
Db	1309	CAAAACGAGGAGATCTGGAAACATATCTTCATGTAGCCATTACATTCACAGAGATTAG	1366

QY	479	ValGlySer-----PropheGluLeuValIaSerGlyAsn-----LysArg	492
Db	1369	CCCGAGAGTAACTTAACCTGTCATATTCATATTCAGAGGCGCATGCATAATTCATGACACAG	1428
QY	493	LeuIysGluLeuSerTyrMetValIaSerArgGlyGlnLeuValaIaValGlyLys---	511
Db	1429	ATCAAAATATTTCATATCCCATATGTAAATAAAGGAAAGATTTTCAAGGTTGGCAGGCAA	1488
QY	512	-----GlnAsnSerThiMetPheSerLeuThrProGluAsnSerTyrThr	526
Db	1489	CCGAGAGAGATGAGCGAAGATCTGGTGACCATGATATCTGCATATCACTCCAGATCTCATC	1548
QY	527	ProIysaIaCysValIleValTyrTyrIleGluAspArgGlyGlnIleIleSerAspVal	546
Db	1549	CTTTCCTCCGGCTTGGCTTGAGCTTACTTACCAAGTGGGAAACAACGAAATTTGGCTGATCTT	1608
QY	547	LeuIysIleProValGln-----LeuValPheLys-----AsnLys	558
Db	1609	GTCGGGGGCGATGTGAAGAGATACCTGCATGGGAACGTTGGTTGAAAGAGACAAATCTA	1668
QY	559	IleIysLeuTyrTyrSerLysValIysAlaGluProSerGluLysValSerLeuArgIle	578
Db	1669	ATACAAATG-----CCAGACACTGCCTCAATGAATCAAAATTTG	1704
QY	579	SerValThrGlnProAspSerIleValGlyIleValaIaValaIaAspLysSerValaIaLeu	598
Db	1705	GAA--GGGATCCAGGTGCTCGGGTGGTCTTGCTGGCTGGCAAAAGCATATATGTT	1761
QY	599	MetAsnAlaSerAsnAspIleThiMetGlyAsnValaIaIleGluLeuGluLeuTyrAsn	618
Db	1762	CTCATATGATAATATAAGATTAGCCCAAGCTTAAGTATAGGACACACAAATAAAGAGTGCAC	1821
QY	619	ThirGlyTyrTyrIleGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeu	638
Db	1822	TTTGGCTGTACAGCTGGCAGCTGGCCAGAAATATCTGGGTGTTTGAAGAGATCTGGACTG	1881
QY	639	TrpValLeuThrAspAlaAsnLeu---ThiLysAsp-----	649
Db	1882	GCTCTGACAAACCAACACTTAATCTTCACACAAACAGATCATGCTGCAAAGTCTCTCAG	1941
QY	649	-----	649
Db	1942	CCTGCAAAATCGSAGGCGTCCGACGTTCTGTTTGTCTGTACAGCAACGACGAAAGCG	2001
QY	649	-----	649
Db	2002	GCAGAAATTTCAGATCAAGACCTCGTAAATGCTGTGAAGATGTCATGATGAGAACCC	2061
QY	649	-----	649
Db	2062	ATGGGGTACACTTGTGAAAGCGGTGCAAAATATACATCCAGAGGAGATGCTTTAAGGCT	2121
QY	650	-----TyrIleAspGlyValTyrAsp---AsnAlaGluTyrAla	661
Db	2122	GCCTTCCCTGATGCTGTCGTTACATCAAGGGGGTCCGAGATGAAGAAACCAACGGGAAGC	2181
QY	662	GluArgPheMet-----GluGluAsnGlu---GlyIleValaIaAspIleHisAspPhe	678
Db	2182	GAGTTGTGTTCTGGCAAGAGATGATATGAGATGGTTTCATACAGATGATGATATCATC	2241
QY	679	SerLeuGlySerSerProHisValaIaArgLysHisPheProGluThrTyrPleIleTrpLeu---	697
Db	2242	TCA-----AGGTCGTCATTTCCCAAGAGTTGGTGTGGCTTAACA	2280
QY	698	-----AspThrAsnMetClyTyrArgIleTyrGlnGluPheGluValaIaThr	712
Db	2281	AAGACCTTGACCGAGAGGCTTAACAGTCAAGGATTTCAACAAACACAACTCTTTTAT	2340
QY	713	ValProAspSerIleThrSerTrpValaIaThrGlyPheValIleSerGluAspLeuGly	732
Db	2341	CTGAGAGATTCATACCAACTGGGTGTGTGGCTGTAAAGCTTTACACCAACCAAGGG	2400
QY	733	LeuGlyLeuThrThrThrProValaGluLeuGlnaIaPheGlnProPheIlePheLeu	752

Db 2401 ATC---TGTGTGGCTGAACCTTATGAAATATAGAGCTAGAAAGCTTCTTATTCATGATCTT 2457  
QY AsnLeuProTyrSerValIleArgGlyGlnGluPheAlaLeuGluIleThrIlePheAsn 772  
Db 2458 CAAATGCCATATTCAGTAGTAGAAGAAATGAGCAGGAGGAGATTCGACGTATTCGACAAAC 2517  
QY TyrLeuIleAspAlaThrGluValIleIleGlySerAspIlePheAspIle 792  
Db 2518 TACGTTAAACGAGATATTATGTATGTGAGATGAACTGTTATCAACCCAGCCTTC----- 2571  
QY LeuMetThrSerSerGluIleAsnAlaThrGlyHisGlnGlnIleuThrLeuValProSer 812  
Db 2572 -----TGCAGTGCCTCCACAAAGAGACAAAGATATCCAGCAGCAGTTCCTCAATTAAGCC 2625  
QY GluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGly-----GluIle 830  
Db 2626 CTGTCTCCAGAGCAGTACCGTTGTGTATGTCCATTAAGAGCAAGATTCGATGATGTT 2685  
QY ProIleThrValThrAlaLeuSerProThrAlaSerAspAlaValThrGlnMetIleu 850  
Db 2686 GAGATTAAAGCAAGGTGCAGAGAACCGTTGTGTGCAGACGGGTGGAGAGAAACTGAAA 2745  
QY ValIleAlaGlnGlyIleGluIleSerTyrSerGlnSerIleLeuLeuAspLeuThrAsp 870  
Db 2746 GTTGATCCTGAAGGGGTACAGAAATCCATGTGACTATTGTTAACTGAGCCAGGGCA 2805  
QY AsnArgLeuGlnSerThr-----LeuIleThrLeuSerPheSerPhe 884  
Db 2806 AAAGAGTGTGGTGAACACACAGCTAGAGATGATCAAAAGCCCAATTAAGATGACAGACTG 2865  
QY ProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIleGlyAspValIleu 904  
Db 2866 CCTGACACAGAAATTGAACCAAGATTATTCATCCAGAGTGACCGCTGTGGCTCAATATT 2925  
QY GlyProSerIleAsnGly-----LeuAlaSerLeuIleArgMetProTyrGlyGlyGly 922  
Db 2926 GAAAACTAAATTGATGAGGAAGTAACTCAACATCTCATTTACCTCTCTGTGGTGTGG 2985  
QY GluGlnAsnMetIleAsnPheAlaProAsnIle-----TyrIleLeuAspTyrIleu 939  
Db 2986 GAGCAAAATATGATCCGATGAGCCGACCACTATTATGCCACTTACTACTGGAACACACA 3045  
QY ThrIleIleIleGlnLeuThrAspAsnLeuIleGlyValAlaLeuSerPheMetArgGln 959  
Db 3046 GAGCAGTGGGAGACTCTCGCATTAATTCGACAGACTGAAGCTGTCAATCAGATCGTACT 3105  
QY GlyTyrGlnArgGlnLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsn 979  
Db 3106 GGTATATGCCAGCAGATGTGTATCAAGAAACAGATCACTTCCTATGCGACTTTTAAAC 3165  
QY TyrAspProSerGlySerThrTyrLeuSerAlaPheValIleuAspGlySerPhe---LeuGln 998  
Db 3166 -----CGTCATCTAGTTCCTTGGCTTAACAGATATGTCGTAAGATCTTGGCATGACT 3219  
QY AlaAspProTyrIleAspIleAspGlnAsnValIleuHisArgThrTyrThrTyrLeu--- 1017  
Db 3220 GCCAAAATGTATGACAGGATTAAGCATATAATCATTTGTGAGAGGTGAGAGGTGCTATTT 3279  
QY LysGlyHisGlnIleSerAsnGlyGluPheTyrAspProGlyArgValIleHisSerGln 1037  
Db 3280 CTGAACAGGACAAACACAGATGAGCGCTTCAAGAAATGCCCCGTGATCTTTCGGAACA 3339  
QY LeuGlnGly-----GlyAsnIleSerProValThrIleuThrAlaTyrIleValThr 1054  
Db 3340 ATGCAAGGAGGAATTCAAAGTGTCTGGAAGAAAGATATTATTAACAGCTTTCATTCGTGT 3399  
QY SerLeuLeuGlyTyrArg-----LysTyrGlnProAsnIleAsp-----Val 1068  
Db 3400 GCGTGTGTGAATCCAAACAACATCTGCAATGACTATGTCATATGCTATGACAGCAGATC 3459  
QY GlnGluSerIleHisPheLeuGluSerGluPheSerArgGlyIleSerAspAsnTyrThr 1088

Db	3460	AAGAAAGCCACAATTAATTATTTCTGAAAAAGATGAGAAA	-CTGCAAAAGCCCTTAACACT	3516
Oy	1089	LeuAlaLeuIleThrTyrrAlaLeuSerSerValGlySerProIlyValAlaValAlaLeu	1108	
Db	3517	ACAGCCCTTCACAGACCTTATGCTTTGGCT	-----GCTGCAGACCACTC	3558
Oy	1109	AsnMetLeuThrTrpAlaGlaGluGlnGlyGlyMetGlnPheTrpValSerSerGlu	1128	
Db	3559	AATGATGACAGGGTACTCATGAGCAGATCAACAGAAAGGATCATTTG	-----	3606
Oy	1129	SerIlyLeuSerAspSerTrpGlnProArgSerLeuAspIleGluValAlaAlaTyrrAla	1148	
Db	3607	-----GAGATATACAAAGCTCACACCCACAACTTGAAGGACCTTCTTAATCC	3655	
Oy	1149	LeuLeuSerHisPheLeuGlnPheGlnThrSerGluGlyIleProIleMetArgTrpLeu	1168	
Db	3655	TTGTTGGCCCTGCGTGAATAAGAAATTTGATGAACTGGTCCCATATGTCAGATGGCTG	3714	
Oy	1169	SerArgGlnArgAsnSerLeuGlyGlyPheAlaSerThrGlnAspTrpThrValAlaLeu	1188	
Db	3715	ACAGATCAGATTTTTATGGGAAAACATATGACAAACCCAAAGACAGATATGGCATTT	3774	
Oy	1189	LysAlaLeuSerGluPheAlaAlaLeuMetAsnThrGluArg	-----ThrAsnIleGlnVal	1207
Db	3775	CAGGTCTCTGTGAATATGAGATTCAAGTTCCTACCCATTAAGACTTAATTAATTAATTT	3834	
Oy	1208	ThrValThrGlyProSerSerProSerProLeu	-----	1218
Db	3895	ACTATTTGAACGCCAGATCCGAGAGTACCTATTAAGTACAGAAATTAATTAAGAAATGCT	3894	
Oy	1219	-----AlaValValGlnPro-----MetAlaValAsnIleSerAlaAsnGly	1222	
Db	3895	CTCCTGGCTCGAGACGTAGAGCCAACTCAACCAAGACATCACTGTGCACAGCATCAGT	3954	
Oy	1233	PheGlyPheAlaIleCysGlnLeuAsnValValTyrrAsnValLysAlaSerGlySerSer	1252	
Db	3955	GATGAGAAAAGCAACATATGACCATTTTGAATTTATTAACGCACGCTGG	4002	
Oy	1253	ArgArgArgArgSerIleGlnAsnGlnGluAlaPheAspLeuAspValAlaValIlyGlu	1272	
Db	4003	CAGAGAAAGCAAAATGTTTGCATATAAA-----TTTCATCTTAATGTTTCTGTGAA	4053	
Oy	1273	AsnIlyAspAspLeuAsnHisValAsp-----LeuAsnValCysThr	1286	
Db	4054	---AACATCCACTGATGATGACATGAGAGCCAAAGGAGCCCTCATGCTCAAGATCTGCACA	4110	
Oy	1287	SerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluValAsnLeuLeuSerGly	1306	
Db	4111	AGGATATCTGGAGAAAGTGAATTCCTACATGACATTAATTAATTTCTATGCTCAGCTG	4170	
Oy	1307	PheMetValProSerGluAlaIleSer-----LeuSerGluThrValIys	1321	
Db	4171	TTTTCTCCCTGATGCTGAGACCTTAACAAGCTTTCTAAAGAGTGCACAGATACATCTCC	4230	
Oy	1322	LysValGluTyrrAspHisGlyIlyLeuAsn-----LeuTyrrLeuAspSer	1336	
Db	4231	AGATATGATGATGACATTAATATATGCTCAGAAATAGTACTGTTATCTTAATTAACAAG	4290	
Oy	1337	ValAsnGluThrGln---PheCysValAsnIleProAlaValArgAsnPheLysValSer	1355	
Db	4291	GTCTCCCACTCTGAAGATGATATGCTCAGCTTAAAGATTCCAGACATTTTGAAGTTGGC	4350	
Oy	1356	AsnThrGlnAspAlaSerValSerIleValAspTyrrTyrrGluProArgGlnAlaVal	1375	
Db	4351	TTTCATTACGCCAGATGATGATCAGAGTGTACAGCTACTACATTAATGAATAATGTAAC	4410	
Oy	1376	ArgSerTyrrAsnSerGluValLysLeuSerSerCysAspLeuCysSerAspValGlnGly	1395	
Db	4411	AAGTTTACCATCCAGATTAAGAAACAGGCTTCTCAATTAAGATATGATATGTAACGTT	4470	
Oy	1396	CysArgProCysGluAspGlyAlaSerGlySerHis-----	1408	
Db	4471	TGCCGATGTGCAGAGAAACCTGTTCTCTGCTCAACATCAGAAAGATGATGATTTCCA	4530	

QY 1409 -----HisSerSerValIlePheIlePheCysPheIleu 1421  
Db 4531 TTACAATTTGAAAAAGCTGGAGACGATGTGATTATGCTACAAAACCAAGCTGCTT 4590  
QY 1422 TyrPheMetGlu 1425  
Db 4591 CGAATAGAGAA 4602

RESULT 9  
US-08-447-411-44  
Sequence 44, Application US/08447411  
Patent No. 5773243  
GENERAL INFORMATION:  
APPLICANT: FRITZINGER, DAVID C.  
APPLICANT: BREDEHORST, REINHARD  
APPLICANT: VOGEL, CARL-WILHELM  
TITLE OF INVENTION: DNA ENCODING COBRA C3, CVF1, AND CVF2  
NUMBER OF SEQUENCES: 81  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
ADDRESS: P.C. Jefferson Davis Highway, Suite 400  
STREET: 1755 S. Jefferson Davis Highway, Suite 400  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/447,411  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/043,747  
FILING DATE: 07-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Oblon, No. 5773243man F.  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 1126-101-0  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 413-3000  
TELEFAX: (703) 413-2220  
TELEX: 248855 OPAT UR  
INFORMATION FOR SEQ ID NO: 44:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5924 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: sig\_peptide  
LOCATION: 4..69  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 70..4929  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 4..4929  
US-08-447-411-44

Alignment Scores:  
Pred. No.: 4,74e-83 Length: 5924  
Score: 885.00 Matches: 359  
Percent Similarity: 40.84% Conservative: 296  
Best Local Similarity: 22.38% Mismatches: 673  
Query Match: 12.04% Indels: 276  
DB: 1 Gaps: 56

US-10-020-095-4 (1-1428) x US-08-447-411-44 (1-5924)  
QY 18 AlaAlaLeuAlaVal-----AlaProGlyProArgPheLeuValThrAla 32  
Db 31 GCTGCTCTATGATTTGGTTTCCAGGGTCTTTCATAGGGGCTCTACACCTCATCACC 90  
QY 33 ProGlyIleIleArgProGlyIleValThrIleGlyValGluLeuLeu---GluHis 51  
Db 91 CTGCTGTTTTCGACACAGACACAGAGCAATTTGGTGGAGCCCATGAGACAGT 150  
QY 52 CysProSerGlnValThrValIleValAla-----GluLeuLeuSerThrAlaSer 67  
Db 151 ACTCCAAACAGGCTTACATCTTTGTCATGATTTTCCAGGAGAGAAACCTTGTTTC 210  
QY 68 AsnLeuThrValSerValIleGluIleGluValIlePheGluIleGlySerPheValThr 87  
Db 211 CAAACAGAGTATGATATGATTCACAGAGAGGATGCTGTCTACT-----CCAACT 261  
QY 88 LeuThrLeuProSerLeuProLeuAsnSerAlaAspGluIleTyrGluLeuArgValThr 107  
Db 262 ATAGAGATTCACGAAAGAGAGAGATGACGACTCCAGGCAAAATCATATGTGTTGTC 321  
QY 108 GlyArgThrGlnAspGluIleLeuPheSerAsnSerThrArgLeuSerPheGluThrIle 127  
Db 322 CAAGTAACTGGTCTCAAGTGAAGATTGAAAGAGTGTCTCTCTTTTCAACGAGTACG 381  
QY 128 ArgIleSerValPheIleGlnThrAspIleValAlaLeuTyrIleProGlyGlnIleValIle 147  
Db 382 -----TTCTGTTTATTCAGACAGATTAAGGATTTATACACAGGCTCTCCAGTAC 435  
QY 148 PheArgIleValThrLeuPheSerAspPheIleValProTyrIleThrSerLeuAsnIleLeu 167  
Db 436 TATCGTGTGTTTCTATGTGATCACAACAGACAGATGAAACAACTGTGATGTTGAG 495  
QY 168 IleValAspProIleSerAsnLeuIleGlnIleThrLeuSerGlnIleSerAspLeuGly 187  
Db 496 TTTCAGACTCCAGAAAGCATTTCTGTCACT-----TCTAATTCACTGACCTTAAC 546  
QY 188 ValIleSerIleThrPheGlnLeuSerSerHisProIleGluIleValProIleSerIle 206  
Db 547 TTCTTC---TGCCCTTACAAATTATCCAGACCTTGTCACTTGGGCACTTGGAGATTTG 603  
QY 207 ---GlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGlnValSerGluTyrVal 225  
Db 604 GCCAAATATGACATTTCCCAAGAGATTTATCGCATTTTGTGATGTCAGAAATATGTC 663  
QY 226 LeuProIlePheGluValThrLeuGlnThrPro-----LeuTyrCysSerMetAsn 242  
Db 664 TTCCCAAGCTTTGAAAGTCCGTGCAACATCAGACAGAAAGTTTTTTTACATTCAGCGCAAT 723  
QY 243 SerIleHisLeuAsnGlnIleThrAlaIleTyrThrTyrGlyIleValProValIleGly 262  
Db 724 ---GAAATTTCCAGCTGTCTATCACTGCAAGGTACTTGATGAGAGAGAGGAGAGGT 780  
QY 263 AspValThrLeuThrPheLeuProLeuSerPheTyrGlyIleValIleValAsnIle----- 280  
Db 781 ---GTGGCTTTCTCTCTTTTGGAGTGAAGAAATAGATGATCTAAAGAGATTTCCAGAC 837  
QY 281 -----ThrIleThrPheIleValIleGlnIleSerAlaAsnPheSerPheAsnAspGlu 298  
Db 838 TCACATCAGCAATTCGATTTATGATGATGAGAGAAACAACTAAAGAGATTAACA 897  
QY 299 MetIleAsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSerSerPro 318  
Db 898 TTCGGTTTCGATTTCCAAATCTCAATGAGCTTGTGGCATATCTGTATGATCT--- 954  
QY 319 GlyProValGluIleLeuThrValThrValThrGluSerValThrGlyIleSerArgAsnVal 338  
Db 955 -----CTAACAGTCATGACAGATATCGAGAGTATGATGATGATGAG 999  
QY 339 SerThrAsnValPhePheIleGlnIleValIleGluPhePheAspTyrThrThr 358  
Db 1000 CAAAGCGCATTCATATGTGGCATCTCCCTATCAGATCCACTTCAAAAACCCCAAA 1059

QY 359 ValLeuLysProSerLeuAsnPhetHraIaThrValLysValThraIaAspGlyAsn 378  
 Db 1060 TATTTCAAGCAGATGCCATGATGAACGCGTGATGTTACCAACCGTATGCTCA 1119  
 QY 379 GlnLeuThrLeuGlnLysArgAsnAsnValIleThyValThrGlnAsnIleTy 398  
 Db 1120 CCAGCT-----GCCCATGTCCAGTGGTATCCAGAGCGCTTT 1155  
 QY 399 ThrGluTyrTrpSerGlySerAsnSerGlyAsnGlnLysMetGlnIleValGlnLysIle 418  
 Db 1156 -----CATTCATGGGAACCATTTGATGATGGACCTGTTAGCTC 1197  
 QY 419 AsnTyrThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGluAspSer 438  
 Db 1198 ATCCGTAACATACCATTTGATGCTCA-----AGCCTACCATTCACCTGTTAAAGCT 1248  
 QY 439 SerGluLeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSer 458  
 Db 1249 AACCATGAGACCTCCCAAGAGAACCCAGGCAACAAAGTCCATGACAGCCCATAGCCTTAC 1308  
 QY 459 LeuPheLysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLys 478  
 Db 1309 CAAACCCAGGAGAGATCTGGAACTATCTTCATGTAGCCATTAATCTACAGAGATTAAAG 1368  
 QY 479 ValGlySer-----ProPheGluLeuValIleSerGlyAsn-----LysArg 492  
 Db 1369 CCCGAGATTAATCTTACTGCTGCAATTTCCATGTGAAGGCCAATGCAATTCACGTGAAGCAG 1428  
 QY 493 LeuLysGlnLeuSerTyrMetValValSerArgGlyGlnLeuValAlaValGlyLys--- 511  
 Db 1429 ATCAATATATTTCACTCACTCATATTTGAATAAGGAGAAATTTTCAAGTGTGGCAGCA 1488  
 QY 512 -----GlnAsnSerThrMetPheSerLeuThrProLysAsnSerTrpThr 526  
 Db 1489 CCAGAGAGATGGAGCAATCTGTGACCAATGATGATCATATCACTCCAGATCTCATC 1548  
 QY 527 ProLysAlaCysValIleValTyrTyrIleGluAspGlyGluIleIleSerAspVal 546  
 Db 1549 CCTTCCTCCGGTTGTGTGCTTACTACCAAGTGGGAAACACGAATTTGTGCTGATTCT 1608  
 QY 547 LeuLysIleProValGln-----LeuValPheLys-----AsnLys 558  
 Db 1609 GTCTGGGTGATGTGAAGATACCTGATGGAACCTGTGTGTGAAGAGACATCTTA 1668  
 QY 559 IleLysLeuTyrTrpSerLysValLysAlaGluProSerGlyLysValSerLeuArgIle 578  
 Db 1669 ATACCAATG-----CCAGAGCTGCAATGAAATCAAAATTG 1704  
 QY 579 SerValThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeu 598  
 Db 1705 GAA---GGGATCCAGGTGCTCGGTTGGTCTGTGGTGGTGGACCAAGCAGATGATGTGT 1761  
 QY 599 MetAsnAlaSerAsnAspIleThrMetGluAsnValValHisGluLeuGluLeuTyrAsn 618  
 Db 1762 CTCATATGATAAATATATAGATTAGCCAAAGTATGGAACACATAGAAAGAGTAC 1821  
 QY 619 ThrGlyTyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeu 638  
 Db 1822 TTTGGCTGTACAGCTGGCAGTGGCCAGAAATACTGGGTGTGTGTGAAGATGCTGAGACGTG 1881  
 QY 639 TrpValLeuThrAspAlaAsnLeu-----ThrLysAsp----- 649  
 Db 1882 GCTGTGACAAACGACACTAATCTCAACACCAACAGAGATCAGCTGCAAAAGTGTCTCAG 1941  
 QY 649 ----- 649  
 Db 1942 CCTGCAATTCGAGAGCGGTGCGAGTTCTGTTTGTCTGCTTGACAGCAACGGAAGCAAGGG 2001  
 QY 649 ----- 649  
 Db 2002 GCAGAAATTCAGAGATCAAGACCTGCGTAATGCTGTGAAGATGTCAATGATGAGAACCCC 2061

QY 649 ----- 649  
 Db 2062 ATGGGGTACACTTGTGAAAAAGCGTGCAAAATATCATCCAGAGAGAGATGCTGTAAAGCT 2121  
 QY 650 -----TyrIleAspGlyValTyrAsp---AsnAlaGluTyrAla 661  
 Db 2122 GCCCTTCCTTGAATGCTGTGCTACATCAAGGGGGTCCGAGATGAAACCAACGGGAGAC 2181  
 QY 662 GluArgPheMet-----GluGluAsnGlu---GlyHisIleValAspIleHisAspPhe 678  
 Db 2182 GAGTTGTTCTGGCAAGAGATGATATATGAAGATGGTTTCAATGACAGATGTGATATCATC 2241  
 QY 679 SerLeuGlySerSerProHisValArgLysHisPheProGluThrTyrIleTrpLeu--- 697  
 Db 2242 TCA-----AGGTGTATTTCCCAAGAGTTGTTGGCTTAACA 2280  
 QY 698 -----AspThrAsnMetGlyTyrArgIleTyrGlnGluPheGluValThr 712  
 Db 2281 AAGCACTTGACCGAGAGCCTAAACAGTCAAGGATTTCAAGCAACAAATGCTTTTAT 2340  
 QY 713 ValProAspSerIleThrSerTrpValAlaThrGlyPheValIleSerGluAspLeuGly 732  
 Db 2341 CTGAGGATTCATCAACAACCTGGGTGTGCTGTGAAGCTTTACCCCAACCAAGGCG 2400  
 QY 733 LeuGlyLeuThrThrProValGluLeuGlnAlaPheGlnProPhePheIlePheLeu 752  
 Db 2401 ATC---TGCTGGCTGAACCTTATGAAATTAAGATCAATGAAGTCTTCTTCATGATCTT 2457  
 QY 753 AsnLeuProTyrSerValIleArgGlyGluGluPheAlaLeuGluIleThrIlePheAsn 772  
 Db 2458 CAAAGCCATATTCAGTGTGAAGATGAGCAGGTGAGATTCGAGCTATTCGACACAA 2517  
 QY 773 TyrLeuLysAspAlaThrGluValLysValIleIleGluLysSerAspLysPheAspIle 792  
 Db 2518 TACGTTAACGAGATATTTATGTGCGAGTGAACCTGTATACACCCAGCCTTC----- 2571  
 QY 793 LeuMetThrSerSerGluIleAsnAlaThrGlyHisGlnGlnThrLeuLeuValProSer 812  
 Db 2572 -----TGACGTGCTTCCCAAAAGACAAAGCAAAATATCCGACACAGTCCCAATTAAGCC 2625  
 QY 813 GluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGly-----GluIle 830  
 Db 2626 CTGTCCCTCAGAGCAGTACCGTTTGATGATGCCATTTAGACCAAGATGATGATGTGT 2685  
 QY 831 ProIleThrValThrAlaLeuSerProThrAlaSerAspAlaValThrGlnMetIleLeu 850  
 Db 2686 GAGATTAAAGCAAGTGTCCAGAGAACGTTGTGTGACGAGTGTGAGGAAGAACTGAAA 2745  
 QY 851 ValLysAlaGluGlyIleGluLysSerTyrSerGlnSerIleLeuLeuAspLeuThrAsp 870  
 Db 2746 GTTGTACCTGAAGGGGTACAGAAATTCATTGTGACTATTTGTTAACTGGACCCAAAGGCA 2805  
 QY 871 AsnArgLeuGlnSerThr-----LeuLysThrLysSerPheSerPhe 884  
 Db 2806 AAAGAGTTGTGTGAACACAGCTAGAGTGAAGCAAGCCCGCAATTTGATGACAGAGTG 2865  
 QY 885 ProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIleGlyAspValLeu 904  
 Db 2866 CTGACACAGAAATGAAACCAAGATTATCAACCAAGTGACCCCTGTGCTCAGATTAT 2925  
 QY 905 GlyProSerIleAsnGly-----LeuAlaSerLeuIleArgMetProTyrGlyCysGly 922  
 Db 2926 GAAAACTCAATGATGAGAAAGTAACTCAACATCTCATATTATCACTCTTGTGCTGTGG 2985  
 QY 923 GluGlnAsnMetIleAsnPheAlaProAsnIle-----TyrIleLeuAspTyrLeu 939  
 Db 2986 GAGCAAAATATGATCCGATGCGCACAGATTAATGACCACTTACCTGACCAACACACA 3045  
 QY 940 ThrLysLysLysGlnLeuThrAspAsnLeuLysGlyLysAlaLeuSerPheMetArgGln 959  
 Db 3046 GAGCAGTGGAGACTCTCGGCAATTAATCGCAGAGATGAGAGCTGTCAATGATCGTGA 3105  
 QY 960 GlyTyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsn 979



```

Db 3106 GGTATCCCGCAGCATGCTGTACAGAAAGCAGATATCTCTATGAGCATTTACAAAC 3165
Qy 980 TyrAspProSerGlySerThrTrpLeuSerAlaPheValLeuAlaGlySerPhe---LeuGlu 998
Db 3166 -----CGTGCATCTAGTCTTGGCTAACACGATATGCGTAAGAGCTTTGCCATGGCT 3219
Qy 999 AlaAspProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrTrpTrpLeu--- 1017
Db 3220 GCCAAATGATGAGCAGCATATGATCATGAATCATTTGTGAGGTGTGAGGTGGCTCAT 3279
Qy 1018 LysGlyHisGlnLysSerAsnGlyGluPheTrpAspProGlyArgValIleHisSerGlu 1037
Db 3280 CTGAACAGGCACAAACAGCATGAGCGCTCAAGAAATATGCCCTGTACCTTCGAAACA 3339
Qy 1038 LeuGlnGly-----GlyAsnLysSerProValThrLeuThrAlaTyrIleValThr 1054
Db 3340 ATGCAGGAGAGAAATTCAGAGTGTGAAAGAGATATTTAAACAGCTTTCATCTGGTT 3399
Qy 1055 SerLeuLeuGlyTyrArg-----LysTyrGlnProAsnIleAsp-----Val 1068
Db 3400 GCGTGTGATGATCCAAACATCTGCATGACTATGTCATAGTCTAGACAGCAGCATC 3459
Qy 1069 GlnGlnSerIleHisPheLeuGlnSerGluPheSerArgIylIleSerAspAsnTyrThr 1088
Db 3460 AAGAAGCCCAAAATATTTACTCAAAAAGTATGAGAAA---CTGCAAGGCGCTTACACT 3516
Qy 1089 LeuAlaLeuIleThrTyrAlaLeuSerSerValGlySerProValAlaLysGlnAlaLeu 1108
Db 3517 ACAGCCCTCACAGCCTTATGCTTGGCT-----GCTGACAGCAACATC 3558
Qy 1109 AsnMetLeuThrTrpArgAlaGlnGlnGlyMetGlnPheTrpValSerSerGlu 1128
Db 3559 AATGATGACAGGATCTCATGCGCAGCATCAACAGGAAGGATCATGCG----- 3606
Qy 1129 SerLysLeuSerAspSerTrpGlnProArgSerLysPheIleGluValAlaAlaTyrAla 1148
Db 3607 -----GAAGATACAAATGCTCAGCACCACCAACATGTAAGGCACTTCATGGCC 3654
Qy 1149 LeuLeuSerHisPheLeuGlnPheGlnThrSerGluGlyIleProIleMetArgTrpLeu 1168
Db 3655 TTGTTGGCCCTCGTGAATAAGTAAGAAATTTGATCAAACTGCTCCATGATGAGTGG 3714
Qy 1169 SerArgGlnArgAsnSerLeuGlyGlyPheAlaSerThrGlnAspThrThrValAlaLeu 1188
Db 3715 ACAATCATGATTTTATGCGGAACATATGACAAACCCAGACAGCATATGCGATTT 3774
Qy 1189 LysAlaLeuSerGluPheAlaAlaLeuMetAsnThrGluArg---ThrAsnIleGlnVal 1207
Db 3775 CAAGCTCTTGCTGAATGATGATTCAGATGCTTACCCATTAAGACTTAACCTTAGATAT 3834
Qy 1208 ThrValThrGlyProSerSerProSerProLeu----- 1218
Db 3835 ACTATTGAACCTGCAGATCGAAGAGTACTTAAAGTACAGAAATTAATTATGAAGATGCT 3894
Qy 1219 -----AlaValAlaGlnPro-----MetAlaValAsnIleSerAlaAsnGly 1232
Db 3895 CTCCTGGCTCGGACAGTGAAGACCAAACTCAACCAAGACATCACTGTACAGCATCAGT 3954
Qy 1233 PheGlyPheAlaIleCysGlnLeuAsnValValTyrAsnValValAlaSerGlySerSer 1252
Db 3955 GATGGAAGAACAAATGACATTTTGACATTTATTAACGACAGCTTG----- 4002
Qy 1253 ArgArgArgArgSerIleGlnAsnGlnGluAlaPheAspLeuAspValAlaValLysGlu 1272
Db 4003 CAGAGAGAGCAAAATGTTTGAATAAA-----TTTCATCTTAATGTTTCTGTTGAA--- 4053
Qy 1273 AsnLysAspAspLeuAsnHisValAsp-----LeuAsnValCysThr 1286
Db 4054 ---AACATCCCTTGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4110
Qy 1287 SerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluValAsnLeuLeuSerGly 1306

```

```

Db 4111 AGGTATCTGGAGAGAGTGTATTTCTACATGACATATATGATTTCTATGCTGACTGCT 4170
Qy 1307 PheMetValProSerGluAlaIleSer-----LeuSerGluThrValLys 1321
Db 4171 TTCTCCCGTATGCTGTAAGACCTTACAGGCTTTCTAAAGAGATGAGACATCATCTCC 4230
Qy 1322 LysValGlnTyrAspHisGlyLysLeuAsn-----LeuTyrLeuAspSer 1336
Db 4231 AGATATCAGAGTTACATATATATGCTCAGAAAGTACCTGTATCTTATCTTAAACAG 4290
Qy 1337 ValAsnGlnThrGln---PheCysValAsnIleProAlaValArgAsnPheLysSer 1355
Db 4291 GTCTCCACCTGTAAGATGATGATCTGCTGATTTAAAGTTCTCAAGCATTTTGAAGTTGGC 4350
Qy 1356 AsnThrGlnAspAlaSerValSerIleValAspTyrTyrGlnProArgArgAlaVal 1375
Db 4351 TTGATTCAGCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4410
Qy 1376 ArgSerTyrAsnSerGluValLysLeuSerSerCysAspLeuCysSerAspValGlnGly 1395
Db 4411 AAGTTCTACCATCCAGATTAAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4470
Qy 1396 CysArgProCysGlnAspGlyAlaSerGlySerHisHis----- 1408
Db 4471 TGCCGATGTGACAGAGAAACCTGCTCTGCTCAACCATGACGAAAGATGATGATTTCCA 4530
Qy 1409 -----HisSerSerValIlePheIlePheCysPheLysLeu 1421
Db 4531 TTACAAATTAAGAAAGCTGCGAGACGAAATGTGATGATGATGATGATGATGATGATGAT 4590
Qy 1422 TyrPheMetGlu 1425
Db 4591 CGAATAGAGAA 4602

RESULT 10
US-08-447-411-75
Sequence 75, Application US/08447411
Patent No. 5773243
GENERAL INFORMATION:
APPLICANT: FRITZINGER, DAVID C.
APPLICANT: BREDEHORST, REINHARD
APPLICANT: VOGEL, CARL-WILHELM
TITLE OF INVENTION: DNA ENCODING COBRA C3, CVF1, AND CVF2
NUMBER OF SEQUENCES: 81
CORRESPONDENCE ADDRESS:
ADDRESSER: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/447,411
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/043,747
FILING DATE: 07-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 5773243man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 1126-101-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 75:

```

```

/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 4138 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: unknown
/ TOPOLOGY: unknown
/ MOLECULE TYPE: DNA (genomic)
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 3..4001
/ US-08-447-411-75

Alignment Scores:
Pred. No.: 6,68e-71 Length: 4138
Score: 768.50 Matches: 306
Percent Similarity: 41.88 Conservative: 212
Best Local Similarity: 24.74 Mismatches: 465
Query Match: 10.46 Indels: 255
DB: 1 Gaps: 46

US-10-020-095-4 (1-1428) x US-08-447-411-75 (1-4138)

QY 349 TyrTlleGluPhePheSeryrThrValLeuLeuProSerLeuAsnPherThra 368
   |||||
DB 66 TATCAGATCTACTTCAAAAACCCCAAAATATTCACAGCCAGAAATGCCATATGAATG 125
   |||||

QY 369 ThrValValValThraGlaAspGlyAsnGlnLeuThrLeuGluGluArgAsnAsn 388
   |||||
DB 126 ACGGTATATGTTACCAAACTGATGCTCACCAAGCTGCC----- 164
   |||||

QY 389 ValValIleThrValThrGlnArgAsnYrThrGluYrThrSerGlySerAsnSerGly 408
   |||||
DB 165 CATGTGCCAGTGTATCAGAGCCATTCATCTGAG-----GGA 203
   |||||

QY 409 AsnGlnMetGluAlaValGlnYrIleAsnYrThrValProGlnSerGlyThrPhe 428
   |||||
DB 204 ACCACTTGTAGTGGAGCTGCTAAGCTCTTCCTGACACACACCAAAATGCTCAA--- 260
   |||||

QY 429 LysIleGluPheProIleLeuGluAspSerSerGluLeuGlnLeuYrIlePheLeu 448
   |||||
DB 261 -----AGCTACCCATCTCTGTAGACATACATGAGACCTCCCAAGAGAACGCCG 314
   |||||

QY 449 GlySerLysSerSerMetAlaValHisSerLeuPheLysSerProSerLysThrTyrIle 468
   |||||
DB 315 GCAATTAAGTCCATGACGACCAAGCTTACCAAAACCCAGAGAGATCTGAAAATATCTT 374
   |||||

QY 469 GlnLeuLysThrArgAspGluAsnIleLysValGlySer-----ProPheGluLeuVal 486
   |||||
DB 375 CATGTAGCCATTACATCTACAGAGATTAAAGCCGGAGATTAACCTTCAATTTCAT 434
   |||||

QY 487 ValSerGlyAsn-----LysArgLeuLysGluLeuSerTyrMetValValSer 502
   |||||
DB 435 GTGAGGGGCAATGCAAAATTCACCTGAACCCAGATCAAAATTTTACATCTTCACTGAAAT 494
   |||||

QY 503 ArgGlyGlnLeuValAlaValGlyLysGlnAsnSer-----Thr 515
   |||||
DB 495 AAAGGGAAGATTTTCAAGGTGGCAGGCAACACAGGAGATGGGAGAAATCGGTGAC 554
   |||||

QY 516 Met---PheSerLeuThrProGluAsnSerTyrThrProLysAlaCysValIleValTyr 534
   |||||
DB 555 ATGAATCTACATATCTCTCCAGAT-----CTCATCTCTTCCGTTTGTGGCTTAC 608
   |||||

QY 535 TyrIleGluAspAspGlyIleIleSerAspValLeuLysIleProValGln----- 552
   |||||
DB 609 TACCAAGTGGAAACATATAATTTGCTGATTTCTGTGGGTGATGTGAAGATACC 668
   |||||

QY 553 -----LeuValPheLys-----AsnLysIleLysLeuTyr 562
   |||||
DB 669 TGCATGGAAAGCTGTGTTGAAAGAGGAGCTTCCAGACATGATGATCAAAATG--- 725
   |||||

QY 563 TrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerValThrGln 582
   |||||
DB 726 -----CCAGAGAGCTGCATGAATGAATCAAAATTTGAA---GGGGAT 761
   |||||

```

```

QY 583 ProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsnAlaSer 602
   |||||
DB 762 CCAAGTGTGGATTTGCTTGTGCTGTGGAACAAAGAGAAATGTTCTCATATATATAA 821
   |||||

QY 603 AsnAspIleThrMetGluAsnValAlaHisGluLeuGluLeuTyrAsnThrGlyTyrTyr 622
   |||||
DB 822 TATAAGATTAGCAAGTAAATATGAGATGAGCAACATAAGAAAGAGTCTGGCTGTACA 881
   |||||

QY 623 LeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuThrValLeuThr 642
   |||||
DB 882 GCTGCACTGGCCAGAAATATCTGGGTGTGTTGAAGATGCTGAGCTGCTGTGCAACC 941
   |||||

QY 643 AspAlaAsnLeuThr----- 647
   |||||
DB 942 AGCATATATCTCAACACCAACAGAGATCAGCTGCAAAAGTGTCTCAGCTGCAATGG 1001
   |||||

QY 647 ----- 647
   |||||
DB 1002 AGGCTGCGCAGTTCTGTTTCTGCTGTGACAGCAACGCAACAAAGCGGACAGTTTCAG 1061
   |||||

QY 648 -----LysAspTyrIleAspGlyValTyrAspAsn----- 657
   |||||
DB 1062 GATCAAGACTGCGTAAATGCTGTGAGATGAGATGAGAGACCCCATGAGGACACT 1121
   |||||

QY 658 AlaGluTyrAlaGluArgPheMetGluGluAsnGluGly----- 670
   |||||
DB 1122 TGTGAAAGCGTGAAATAATACATCCAGAGGAGATGCTTGAAGCTGCTCTCGAA 1181
   |||||

QY 671 -----HisIleValAspIleHisAsp----- 677
   |||||
DB 1182 TGTCTGTACTATCATCAAGAGATCCAGATATCAATAAAGCGAGAGCGAGTGTGTTCTG 1241
   |||||

QY 678 -----PheSerLeuGlySerSerProHisValArgLys 688
   |||||
DB 1242 GCAGAAAGTATTTGAAGATGATTTATTTGAGAGAGGTAACATCACTCA---AGGCT 1298
   |||||

QY 689 HisPheProGluThrTrpIleTrpLeu-----AspThr 699
   |||||
DB 1299 GATTTTCTGAGAGTGTGTTGCTTATGAGAGAGCTGTCTGAACATCTTACAGCTAAA 1358
   |||||

QY 700 AsnMetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIleThrSer 719
   |||||
DB 1359 GCGATTTCAGACAGATATGTAACCTTTAT-----CTGAGGAGATTCATCAACACC 1409
   |||||

QY 720 TrpValAlaThrGlyPheValIleSerGluAspLeuGlyLeuThrThrThr--- 738
   |||||
DB 1410 TGG-----GAGTTGCTGGCTGTGGGCTTTCAACCCACCAAA 1445
   |||||

QY 739 -----ProValGluLeuGlnAlaPheGlnProPhePheIlePheLeu 752
   |||||
DB 1446 GGGATCTGTGTGGCTGAACCTTATGAAATTAACAGTCATGAAAGACTTTTCAATTGATCTT 1505
   |||||

QY 753 AsnLeuProTyrSerValIleArgGlyGlnGluPheAlaLeuGluIleThrIlePheAsn 772
   |||||
DB 1506 CAATGCGCGTATTCAGTGAAGAATGACGAGGGAATAATTGCGATGTTTGTGAACAC 1565
   |||||

QY 773 TyrLeuLysAspAlaThrGluValLysValIleIleGluLysSerAspLysPheAspIle 792
   |||||
DB 1566 TACGTGACAAAGATATTATTAAGATGGAACGTATTAACAGCCGCTTTCGACGT 1625
   |||||

QY 793 LeuMetThrSerSerGluIleAsnAlaThrGlyHisGlnGlnThrLeuLeuValProSer 812
   |||||
DB 1626 GCTTCACAGAAAGTCA-----AGATTCAGAGCAAGTTCGCAATTAAGCC 1673
   |||||

QY 813 GluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGly-----GluIle 830
   |||||
DB 1674 CTGTCTCCAGAGGAGTATCTGTTGTATAGCCCATTAAGCAAGATGATGATGATTT 1733
   |||||

QY 831 ProIleThrValThrAlaLeuSerProThrAlaSerAspAlaValThrGlnMetIleLeu 850
   |||||
DB 1734 GAGGTTAACAGAAAGTGTCCAGAGAGATTGATGTCAGATGTGTGAAGAAAGAAATCTGAA 1793
   |||||

QY 851 ValLysAlaGluGlyIleGluLysSerTyrSerGlnSerIleLeuLeuAspLeuThrAsp 870
   |||||

```

```

Db      1794 GTTGTACCTGAAGGGAATGAAAGATTTGTTACTATTATTTGAATGAGCCACATATCA 1853
Qy      AsnArgLeuGlnSerThr-----LeuYsThrLeuSerPheSerPhe 884
Db      1854 AAAGCAATTTGGTGAACACAGGTAGATTTGGTCAAAAGCCAAATTAATTTAAAGACAGGCTT 1913
Qy      ProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIleGlyAspValLeu 904
Db      1914 CCTGATACGGAATATGAAGAACAGATTTACTATTACAGGTATCCGTGCTCAGACATATT 1973
Qy      GlyProSerIleAsnGly-----LeuAlaSerLeuIleArgMetProTyrGlyCysGly 922
Db      1974 GAAACCTCAATTTGATGGAAGTAACTCAACCATCTCATATTCATCTCTTTGGCTGTGGG 2033
Qy      GluGlnAsnMetIleAsnPhe---AlaProAsnIle-----TyrIleLeuAspTyrLeu 939
Db      2034 GAGCAAAATATATGATCCGATGACCTGACACCAAGTTATTCGCCACCTTACTTACGACCA 2093
Qy      ThrIlySlySlyGlnLeuThrAspAsnLeuSlyGlyValAlaLeuSerPheMetArgGln 959
Db      2094 CAGAGTGGAGACTCTCGGATAAATCGCAGACGTGAAGCTGTCAATCAGATCATGACT 2153
Qy      GlyTyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsn 979
Db      2154 GGTATATGCCACAGCTTGTTGTACAAAGAACACAGCATTCCTTATGACGATTTTACAAAC 2213
Qy      TyrAspProSerGlySerThrTyrLeuSerAlaPheValLeuArgCysPhe---LeuGlu 998
Db      2214 -----AGTGCATCTAGTTCTTGGCTTAAACGATATGTTGTAATAAATCTTTGCCCTTGGCT 2267
Qy      AlaAspProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTyrLeu--- 1017
Db      2268 GCCAAATATGTAAGATTAACATTAACGTAATTCGTTGTGAGATATAGGTGCTGATTT 2327
Qy      LysGlyHisGlnLysSerAsnGlyGlnPheThrAspProGlyArgValIleHisSerGlu 1037
Db      2328 CTGAACAGGCAACGACAGATGAGTGTTCAGAGAAACGCCCTGTACTTTTGGAAACA 2387
Qy      LeuGlnGlyIleAsnLysSerProVal-----ThrLeuThrAlaTyrIleValThr 1054
Db      2388 ATGCAGGAGGACATTCAGAGTCTGACACCAAGAGATTTTAAACAGCTTATCTGCTT 2447
Qy      SerLeuLeuGlyTyrArgLysTyrGlnProAsnIleAspValGlnLysSerIleHisPhe 1074
Db      2448 GCGTTGTTGGAATCCAGATCAATC-----TGCATGCAATATATCAATATT 2492
Qy      LeuGlySerGluPheSerArgGly-----IleSer 1084
Db      2493 CTAAACAGCAGCAGATCAATGAGCCACAGATTTTACTCAAAAAGTATGAGAAACCTGCA 2552
Qy      AspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGlySerProValAla 1104
Db      2553 AGGCCCTTACATACAGCCCTTACAGCCCTTATGCTTTGGCT-----GCT 2594
Qy      LysGluAlaLeuAsn-Met-----LeuThrTyrArgLagLuglnGlyMetGly 1122
Db      2595 GCAGAACACATCTATGATGACAGGTTACTATGCGAGATCAACAGAGAA----- 2646
Qy      nPheTyrValSerSerGluSerLysLeuSerAspSerTyrGlnProArgSerLeuAspIle 1142
Db      2647 -----ATCGTTGGAGAGAACTTAACGCCACACACCATATAC-AT 2683
Qy      eGluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnIleGlnIleThrSerGlyIle 1162
Db      2684 TGAAGGCACTTCCATATGCTTGTGGCCCTCTGTAAGAAATTTGTTGGGCGCG 2743
Qy      ePrlIleLeuArgTyrLeuSerArgGlnArgAsnSerLeuGlyGlyPheAlaSerThrGly 1182
Db      2744 TCTGTATCTCCAAATGCTGATGATCAGCAATATTTATGGGGAACATATGCAAAACCA 2803
Qy      nAspThrThrValAlaLeuLysGluPheAlaLeuMetAsnThrGluArg 1202

```

```

Db      2804 AGCAACGTTATGATGTTTCAAGCTCTTGCTGATATATGAGATTCAGATCCCTATAA 2863
Qy      g---ThrAsnIleGluValThrValThrGlyProSerSerProSerPro----- 1217
Db      2864 GCACTTAACTTATGATATTTACTATGAACTCCAGATCGAAGATGAAAGTACTTAAAGTACAG 2923
Qy      1218 -----LeuAlaValAlaGlnPrometAlaValAsn----- 1227
Db      2924 AATTATTTATGAAATATGCTCTCTGCTGCTGACAGATGAGACCAACTCAACCAAGACTT 2983
Qy      1228 -----IleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValTyrAsnVa 1246
Db      2984 CACTGTGTCAGATCAGGTATGTAAGAAAGCAACATGACCATTTTACGCTTATATATGC 3043
Qy      1246 llyAspLysGlySerSerArgArgArgSerIleGlnAsnGlnIleAlaPheAspLe 1266
Db      3044 ACAATTC-----AGGAGAGATGCAAAATGTTTGCACAAA-----TTCCATCT 3085
Qy      1266 uAspValAlaVal-----LysGluAsnLysAspAspLeuAsnH1 1279
Db      3086 TGATGTTTCTGTGAAAACGTCACAGTTGAACTTAAAGAGCAAGGCAAGGAGC 3145
Qy      1279 sValAspLeuAsnValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMe 1299
Db      3146 CTTCAAGCTCAAAATCTGACCTAGATCTGAGAGAGTTGATTTTCAATGACCAATAT 3205
Qy      1299 tGluValAsnLeuLeuSerGlyPheMetValProSerGluAlaIleSer---LeuSerGly 1318
Db      3206 TGATGTTTCTATAGCTGACTGCTTTTGTCTGCTGATCTGAAAGACTTTCAGGCTTTCTAA 3265
Qy      1318 urhValSlyLys-----ValGluTyrAspHis-----GlyLys 1329
Db      3266 AGAGTGCAGACATATATCTCCATGTTTGAATTAACATATATATGCTCAGAAAGGAC 3325
Qy      1329 sLeuAsnLeuTyrLeuAspSerValAsnGluThrGln---PheCysValAsnIleProAl 1348
Db      3326 TGTATTCATTTTACTTATGACAGAGTCTCCACTGTAAGATGAATGCTTGACCTTTAAGAT 3385
Qy      1348 aValAspAsnPheLysValSerAsnThrGlnAspAlaSerValSerIleValAspTyrTyr 1368
Db      3386 TCTCAACATTTTGAATTTGGCTTCACTTACGACAGATCAGTCAAGGTATGACACTTCTA 3445
Qy      1368 rGluProArgArgGlnAlaValArgSerTyrAsnSerGluValLysLeuSerSerCysAs 1388
Db      3446 CAATCTGATGAAAGAAATGTAACAAGATCTTCAATCCAGATGAGAGAACAGAGCTTCTCA 3505
Qy      1388 pleuCysSerAspValGlnGlyCysArgProCysGluAspGlyAlaSer 1404
Db      3506 TAAGATATGTTGTTGTAACGTTTGGCGATGTGCAAGAAACCTGTTCC 3554

RESULT 11
US-08-662-227-33
Sequence 33, Application US/08662227
Patent No. 5922320
GENERAL INFORMATION:
APPLICANT: VOGEL, CARL-WILHELM
APPLICANT: BREDEHORST, REINHORST
APPLICANT: KOCK, MICHAEL
APPLICANT: FRITZINGER, DAVID
TITLE OF INVENTION: RECOMBINANT PROCVF
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBION, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESSER: P.C. JEFFERSON DAVIS HIGHWAY
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

```

SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/662,227  
 FILING DATE: 14-JUN-1996  
 CLASSIFICATION: 530  
 ATTORNEY/AGENT INFORMATION:  
 NAME: OBLON, NORMAN F.  
 REGISTRATION NUMBER: 24,618  
 REFERENCE/DOCKET NUMBER: 1126-0107-0X  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 703-413-3000  
 TELEFAX: 703-413-2220  
 INFORMATION FOR SEQ ID NO: 33:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 4138 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 US-08-662-227-33

## Alignment Scores:

Pred. No.:	6 686-71	Length:	4138
Score:	768.50	Matches:	306
Percent Similarity:	41.88%	Conservative:	212
Best Local Similarity:	24.74%	Mismatches:	465
Query Match:	10.46%	Indels:	255
DB:	2	Gaps:	46

US-10-020-095-4 (1-1428) x US-08-662-227-33 (1-4138)

```

QY 349 TTTTLEGLUPHEPHEAPRYRTHRTHRVALLLEULYPSROSERLEUASNPHETHRALA 368
DB 66 TATGATGATCTACTTCAAAAACCCCAAAATATTTCAGCCAGATGCGATATGAACTG 125
QY 369 TTTTLLYSLVATHTHRGALAAAPGLYASNGINLEUTHRLGUGLUGARGAENAN 388
DB 126 ACGGTGTATGTTTACCAACCTGATGCTCACCCAGCTGCC----- 164
QY 389 VALVALILETHRVATHRGINARGAENRYRTHGULYRTRTPSERGLYSERASNSERGLY 408
DB 165 CATGTGCCAGTGTATCAGAGGCCATTCATTCGAG-----GGA 203
QY 409 AANGINLYMEGLUALAVAGLINSILEASNYRTHRVALLPROGINSERGLYTHRPH 428
DB 204 ACCACTTGAAGTATGAGCTGTAAGCTCTTCTTAACACACACCAAAATGCTCA--- 260
QY 429 LYSILLEGUPHEPROLLEUGLUAASPSERGLULEUGLULEULYSALATYRPHLEU 448
DB 261 -----AGCTACCCGATCACTGTTAGACTAACCATGAGAGACCTCCAGAGAGCCG 314
QY 449 GLYSERLYSERSERMETALAVAIHISERLEUPHELYSERPROSERLYSHRYTLE 468
DB 315 GCAATTAAGTCCATGACACGCCACAGCTTACCAAAACCCAGAGAGATCTGAAAATATCT 374
QY 469 GINLEULYETHRARGSPGLUASNILELYSVALIGLYSER-----PROPHGILULEUVAL 486
DB 375 CATGTAGCCATTAATCTACAGAGATTAAGCCCGAGATTAATCTACTGCAATTTCAAT 434
QY 487 VALISERGLYASN-----LYSARGLEULYSGLULEUSERTYRMEVALVALSER 502
DB 435 GTGAGGGGCAATGCAAAATTCACCTGACACAGATCAAAATATTTTCAATACCTACATG 494
QY 503 ARGGLYGLINLEUVALAVAGLYLYSGINASNSE-----THR 515
DB 495 AAAGGGAAGATTTTCAAGGTGGCAGGCAACACAGGAGAGATGGGAGAAATCTGATGAC 554
QY 516 MET-----PHESERLEUTHRPROGLUAENSERTPTTHRPROLYSALACYSTVALILEVAL 534
DB 555 ATGAATCTACATATCTCACTCAAGT-----CTCATTCCTTCTCCGCGTTTGCGCTTAC 608
QY 535 TTTTLEGLUASAPGLYGLULIILESERASPVALLLEULYSLLEPROVALGIN----- 552
  
```

```

DB 609 TACCAATGGGAAACAATGAATTTGTGCTGATTTCTGTGGGTGAGATGAGATACC 668
QY 553 -----LEUVALPHELY-----ASNLYILELYSEUTHYR 562
DB 669 TGCATGGGAGACGTTGTGTGAAAGAGAGGACTTCCAGAGACATGAAATGCAAAATG--- 725
QY 563 TRPSERYSLVALYSLAGLUPROSERGLULYSERLEUARGILESERVALTHRGIN 582
DB 726 -----CCAGAGCTGCAATGAATAATCAATTTGAA-----GGGAT 761
QY 583 PROAPSERILEVALGLYILEVALAVALSPLYSSERVALASNLEUMETASNAASER 602
DB 762 CCAGGTGCTTGATGTTGCTTGTGCTGTGGCAAGAGATATGTTCTCATATATAA 821
QY 603 ASNAAPLLETHMETGLUASNVALVAHISGLIULEUGLULEUTHYRASNTHGLYTRYR 622
DB 822 TATTAAGATTAGCCAAAGCTAAGATATGGACACAAATGAAAGAGTACTTGGCTGTACA 881
QY 623 LEUGLYMETPHEMETASNSESPHEALAPHEGLINUCYSGLYLEUTHRYVALLEUTHR 642
DB 882 GCTGCAGTGGCCAGAAATATCTGGGTGTGTGTAAGATGCTGCACTGGCTGTGCAAC 941
QY 643 ASPALASNLEUTH----- 647
DB 942 AGCACTTAATCTCAACACCAACAGAGATCAGTCAAAAGTCTCAGCTTCAATCGG 1001
QY 647 ----- 647
DB 1002 AGCGCTCGAGTTCGTTTGTCTGCTTGAACAGCAACGAGCAAGCGGACAGTTTCAG 1061
QY 648 -----LYASPTYLEASPTYLEASPTYLEASPTYLEASPTYLEASPTYLEASPT 657
DB 1062 GATCAAGACCTGCGTAATATGCTGTGAAGATGTCATGACATGAGAACCCATGGGCACT 1121
QY 658 ALAGLUTYRALAGLARGPHEMETGLULUBSNGULY----- 670
DB 1122 TGTGAAAAGCCGTGAATAATATCATCCAGAGGAGATGTTGTAAGTCTGCTTCCGAA 1181
QY 671 -----HISLLEVALASPLLEHISAP----- 677
DB 1182 TGTCTGTCACTACATCAAGAGATCCAGATGCAATTAACGGAGAGAGTGTGTTCTG 1241
QY 678 -----PHESERLEUGLYSERSEPHISVALARYLS 688
DB 1242 GCAAGAAGTATTTGAAGATGATTTATTTGGAGAAGTAACTACATCACTCA---AGCTCT 1298
QY 689 HISPEPROGLUTHTRTPTLLETPLEU-----ASPTHR 699
DB 1299 GATTTTCTGAGAGTGTGTGCTTAATGAGCAGCTGTGAAACATCTTAACAGTAA 1358
QY 700 ASNMETGLYTRYARGILETYRGINLUPHEGLUVALTHRVALLPROAPSERILETHSER 719
DB 1359 GGGATTTCAAGCAAGATGATCTTTTAT-----CTGAGGAGATTCATCAAC 1409
QY 720 TRPVALLATHTHGLYPHEVALILESERGLUASPLEUGLYLEUGLULEUTHRTHR--- 738
DB 1410 TGG-----GAGTTCGCTGCTGTGGGCTTTCACCCACCA 1445
QY 739 -----PROVALGULEUGLUALAPHEGINPROPHPEHLEU 752
DB 1446 GGGATCTGTGTGCTGAGCACTTAATGAATAACAGCATGAAGAAGACTTCTTATGATCT 1505
QY 753 ASNLEUPROTYRSEVALILEARGLYGLUGLUPHEALALEUGLULIETHRILEPHEAN 772
DB 1506 CAACGTCCGTATTCAGTGTGAGATAGAGCGTGAATAATTCGAGCTGTTTGTACAC 1565
QY 773 TTYRLEULYASPALATHRGULVALYSVALILELEGLIULYSERASPLYSEAPLYSE 792
DB 1566 TACGCTGACAGAGATTTATATGATGAGTGAACCTTATACAGCCGCTTCTGCACT 1625
QY 793 LEUMETHTRSESERGLULIASNALATHRGLYHISGLINGINTHRLEUVALPROSER 812
DB 1626 GCTTCCACAGAAAGTCA-----AGATCCGAGAGAGCTGTCATTTAAAGCC 1673
  
```

QY 813 GluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGly-----GluIle 830  
 DB 1674 CTGCTCCAGGAGGAGTATCGTTGTGTATGATCCATTAGACAGAGATTGCATGAGTGT 1733  
 QY 831 ProIleThrValThrAlaLeuSerProThrAlaSerAlaValThrGlnMetIleLeu 850  
 DB 1734 GAGGTTCACAGCAAGTGTCCAGGAGAGTGTGATGTGATGTGTGAGAGAAACCTGAAA 1793  
 QY 851 ValIysAlaGluGlyIleGluLeuSerTyrSerGlnSerIleLeuLeuAspLeuThrAsp 870  
 DB 1794 GTTGACTGTAAGGGAGATGAGAAAGATGTTGTTACTATTATTGAACTGGACCCACATACA 1853  
 QY 871 AsnArgLeuGlnSerThr-----LeuLysThrLeuSerPheSerPhe 884  
 DB 1854 AAAGGAATGTGGACACACAGGTAGAAATTGTCACAAGCCATAAATTAAATGACAGGGTT 1913  
 QY 885 ProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIleGlyAspValLeu 904  
 DB 1914 CCTGATACGGAAATAGAAACCAAGATTACTATCTATTCAGATGATCCTGTGGCTCAGACTATT 1973  
 QY 905 GlyProSerIleAsnGly-----LeuAlaSerLeuIleArgMetProTyrGlyCysGly 922  
 DB 1974 GAAAACTCAATTGATGAGAGTAACTCAACCATCTCATTTACTCTCTTTGGCTGGGG 2033  
 QY 923 GluGlnAsnMetIleAsnPhe---AlaProAsnIle-----TyrIleLeuAspTyrLeu 939  
 DB 2034 GAGCAAAATATGATCCGATGACGTGACACCACTTATTCGACCTTACTCTGACACCA 2093  
 QY 940 ThrIysIysGlnLeuThrAspAsnLeuLysGluValAlaLeuSerPheMetArgGln 959  
 DB 2094 CAGACGTGGAGACTCTCGCATTAATCGCAGAGCTGCAAGCTGCATCAATCATGCTACT 2153  
 QY 960 GlyTyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsn 979  
 DB 2154 GGTATGCTCCAGACGTTGTGTACAGAAACAGACCATCTCTATGACGACATTTACAAAC 2213  
 QY 980 TyrAspProSerGlySerThrTyrLeuSerIleAlaPheValLeuArgCysPhe---LeuGln 998  
 DB 2214 -----ACTGCATCTAGTCTTGCTGCTACACATATGTGTAAAAATTTTGGCTGGCT 2267  
 QY 999 AlaAspProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTyrLeu--- 1017  
 DB 2268 GCCAAATTTGTAAGAACATTAACCATGAAATCGTTGTGAGGATGAGGTGCTGAT 2327  
 QY 1018 LysGlyHisGlnLysSerAsnGlyGluPheThrAspProGlyArgValIleHisSerGlu 1037  
 DB 2328 CTGAACAGGCACAGACAGATGAGTGTTCAGAGAAAAGCCCTGTACTTTTGGACA 2387  
 QY 1038 LeuGlnGlyIleAsnLysSerProVal-----ThrLeuThrAlaTyrIleValThr 1054  
 DB 2388 ATCGAGGAGGACATTCMAAGTGTGACACCAAGAGATCTTTAACGCTTCACTGTGTT 2447  
 QY 1055 SerLeuLeuGlyTyrArgLysTyrGlnProAsnIleAspValGlnLysSerIleHisPhe 1074  
 DB 2448 GCGTTGTGGAATCCAGATCAATC-----TCGAATGATATATACAAAT 2492  
 QY 1075 LeuGlnSerGluPheSerArgGly-----IleSer 1084  
 DB 2493 CTAGACAGACATCAGTAAGGCCACAGATTATTACTCAAAAAGTATGAGAACTGCA 2552  
 QY 1085 AspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerValGlySerProLysAla 1104  
 DB 2553 AGGCTTACACTACAGCCCTCAGACCCATATGCTTTGGCT-----GCT 2594  
 QY 1105 LysGlnAlaLeuAsn-Met-----LeuThrTyrArgAlaGlnGlnGlyMetGln 1122  
 DB 2595 GCAGAACGACATCAATGATGACAGGGTACTCATGGACATCAACGAGAAAGA----- 2646  
 QY 1122 nPheTyrValSerSerGlnSerLysLeuSerAspSerTyrGlnProArgSerLeuAspIle 1142  
 DB 2647 -----ATCGTTGGAGAAACCTTAAGCCCAACCCATAC-AT 2683

QY 1142 eGluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGluGlyIle 1162  
 DB 2684 TGAAGGACATCTCCATCTGTTGTGGCTCGTGAATAAGAAATTTGTTAGGCGCG 2743  
 QY 1162 eProIleMetArgTyrLeuSerArgGlnArgAsnSerLeuGlyGlyPheAlaSerThrGln 1182  
 DB 2744 TCCGTATGTCAAATGGCTGTATGATGACCAATATTATGGGGGAACATATGACAAACCCA 2803  
 QY 1182 nAspThrThrValAlaLeuLysAlaLeuSerGluPheAlaAlaLeuMetAsnThrGluArg 1202  
 DB 2804 AGCAACAGTATATGATGTTTCAAGCTCTGCTGAATATGAAATTCAGATGCTTACCCATA 2863  
 QY 1202 g---ThrAsnIleGlnValThrValThrGlyProSerSerProSerPro----- 1217  
 DB 2864 GCACTTAACCTTATGATATCTATCTATTAAGTCCAGATCGAAGTACTTAAAGTACAG 2923  
 QY 1218 -----LeuAlaValAlaGlnProMetAlaValAsn----- 1227  
 DB 2924 AATTAAATTAGAAATGCTCTCTGCTCAGACAGTACAGACCAAACTCAACGACATT 2983  
 QY 1228 -----IleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValValTyrAsnVal 1246  
 DB 2984 CACTGTCTCAGCATCAGTGTATGAGAAAGCAAAATGACATTTTGAAGCTCTTATATGC 3043  
 QY 1246 LysAlaSerGlySerSerArgArgArgSerIleGlnAsnGlnAlaPheAspLe 1266  
 DB 3044 ACAATTS-----AGGAGAGATGCAATGTTTGCACAA-----TTCCATCT 3085  
 QY 1266 nAspValAlaVal-----LysGluAsnLysAspAspLeuAsnHis 1279  
 DB 3086 TGAATGTTTCGTGGAACCGTCCAGTTGAACCTTAAAGAGCAAGGAGCCAAAGGAGC 3145  
 QY 1279 sValAspLeuAsnValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMet 1299  
 DB 3146 CTCACAGCTAAATCTGCACCTAGATCTCGGAGAAAGTTATCTTACATGACAAATAT 3205  
 QY 1299 tGluValAsnLeuLeuSerGlyPheMetValProSerGluAlaIleSer---LeuSerGln 1318  
 DB 3206 TGAATGTTTATGCTGACTGCTGTTTGGCCCTGATATGAAAGCCTTACGAGGCTTGTAA 3265  
 QY 1318 uThrValLysLys-----ValGluTyrAspHis-----GlyLys 1329  
 DB 3266 AGGAGTGCAGATATATCTCCATGTTTGAATTAACAAATATATGCTCAGAAAGGAAAC 3325  
 QY 1329 sLeuAsnLeuTyrLeuAspSerValAsnGluThrGln---PheCysValAsnIleProAl 1348  
 DB 3326 TGTATATATTTACTTACAGAGGTTCTCCACTGGAAGATGATGCTGCATTTAAGAT 3385  
 QY 1348 aValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSerIleValAspTyr 1368  
 DB 3386 TCTCAACATTTTGAAGTTGCTTCAATTCAGCCAGATCAGTCMAAGGTGTACAGCTACTA 3445  
 QY 1368 tGluProArgArgGlnAlaValArgSerTyrAsnSerGluValLysLeuSerSerCysAs 1388  
 DB 3446 CAATCTAGATGAAAAAAGTACCAAGATCTACATCCAGATGAAAGCAACAGGCTTCTCAA 3505  
 QY 1388 pLeuCysSerAspValGlnGlyCysArgProCysGluAspGlyAlaSer 1404  
 DB 3506 TAAGATATGTGTTGTGTAACGTTTGCCGATGTGCAGAAAGAAACCTGTTCC 3554

RESULT 12  
 US-09-017-947-33  
 ; Sequence 33, Application US/09017947  
 ; Patent No. 6303754  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VOGEL, CARL-WILHELM  
 ; APPLICANT: BREDEHORST, REINHORST  
 ; APPLICANT: KOCK, MICHAEL  
 ; APPLICANT: FRITZINGER, DAVID  
 ; TITLE OF INVENTION: RECOMBINANT PROCVF  
 ; NUMBER OF SEQUENCES: 39  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: OBION, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,



QY 416 -----Glnlyslleasn-----TyrThValProGln----- 424  
 Db 16768 CGATGCCGCAACATACGATACGCTTCGCTTATACCGACCTTCAGATGCTGAGCC 16709  
 QY 425 -----SerGlyThrPheLysIleGluPheProIleLe 435  
 Db 16708 TGCCTCCGGCCGATGGGCGAAGTTCCTTCGGGACAGACGATGCGTTTGGCGAG 16649  
 QY 435 u-----GluAspSerSerGluLeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSe 453  
 Db 16648 TGGCGGAGAGGATTCGCACTTGTTCAGACAGAGCTTATCTCTCCGTCGGAAGATCG 16589  
 QY 453 rmeAlaValHisSerLeuPheLysSerProSerLys-----ThrTyrIleGlnLe 470  
 Db 16588 TCAG-----TTGTGCAATCCTCTCCCAAGGCTTCGACATATGTGGCGGA 16544  
 QY 470 uLysThrArgAspGluAsn-----IleLysValGlySerPro----- 482  
 Db 16543 GGAGAAATACAGACAGTACCGCCCGCCGATCCTGGTGGTCTACTCCCAAGATGC 16484  
 QY 483 -----PheGluLeuValValSerGlyAsnLysValGleuLy 494  
 Db 16483 GGATGCTCAGAGTTGATCTGTTCTACGACCTGACTCAGAGCGC-----CGATTCTAT 16430  
 QY 494 sglLysSerTyrMetValValSerArgGlyGlnLeuValAlaValGlyLysGlnAnse 514  
 Db 16429 CGAAGAGAAATGATCGCTTCGCGCGCGGATGATCGTCAAGATA-----GT 16382  
 QY 514 rThMetPheSerLeuThrProGluAsnSerTyrThrProLysAlaCysValIleValIly 534  
 Db 16381 GCCGATGCTCCGCGACGCTCG-----TTGCGCGAGATGATGATGATGTTCTCT 16334  
 QY 534 rTyrIleGluAspArgIleGluIleIleSerAspValLeuLysIleProValGlnLeuVa 554  
 Db 16333 CTATTATGATATAGATGAGGAGGCTGTATATGCAAGCTAGATCTCGAAGCCAGATGCC 16274  
 QY 554 lPheLysAsnLysIleLysLeuTyrTrpSer-----LysValIlyAsnIleGluProSerG 572  
 Db 16273 C--AAGGAGAGATTAAGATGCTGTGAGTACTTTCGCGACAGCTCCGACCTGGTGA 16217  
 QY 572 uLysValSerLeuArgIleSerValThrGlnProAspSerIleValGlyIleValAlaVa 592  
 Db 16216 GAAAGAGACATGAGATCTCCCTTCGCGGATCGGAG-----GGCAA 16175  
 QY 592 lAspLysSerValAsnLeuMetAsnAla--SerAsnAspIleThrMetGluAsnValVa 611  
 Db 16174 ACCGCTCTCATACCATGATGCGCTGTGATGATGATGATGCTCTGTATAGATCGT 16115  
 QY 611 lHis----- 612  
 Db 16114 GCACAGTCTGTTTCTCTCCGATTCCTCTTCTTTCGCGACTCCGCGCTCTTGTGT 16055  
 QY 613 -GluLeuGlnLeuTyrThrArgIleTyrTyrLeuGlyMetPheMetAsnSer----- 629  
 Db 16054 TAGGTTCCTTCCGATACGAGGATTCAGATGCTGGGGTTCGTGGAGATTCCTTACGT 15995  
 QY 630 -----PheAlaValPheGlnGlnLysCysGlyLeuTrpValLe 641  
 Db 15994 GCAAGTCGGGTTCAAACACCTGTTTCAAAACTCCCAACTGCG--CTGTGGTCTGCG 15938  
 QY 641 uThrAspAlaAsn--LeuThrLysAspTyrIleAspGlyValIlyTrpAsp----- 656  
 Db 15937 TACCGTTATAGATATGTGTTCGGATTAT--TATTATGATGAGCCGATGAT 15887  
 QY 657 -----AsnAlaGluTyrAlaGluLysPhe--MetGlnGluLys 668  
 Db 15886 ATTGTGGGATATGAGATTCGGAAGCCAAATCGGTGATGCTTCCGAGTGGAGAGAA 15827  
 QY 668 nglGluIHisIleValAspIleHisAspPheSerLeuGlySer----- 682  
 Db 15826 AGCTGCTCGGCGAGATGCGTAATGAAAGAGAAACCCGAAATCCGAGATTCGA 15767  
 QY 683 -SerProHisValArgLysHisPheProGluThrTripletTrp-----LeuAspThr 699

Db 15766 AGAATCTCCGTCGATCAACATTCCTGAGACTGATCTTTCAGCCGCGATTCGTAC 15707  
 QY 699 rAsnMetGlyTyrArgIleTyrGlnIlePheGluValThrValProAspSerIleThrSe 719  
 Db 15706 CGATGAACGGGGGAGATTTCTTGATCTTTC-----ACTCTCCGAGACATCGACGCG 15653  
 QY 719 rTrpValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrTrp 739  
 Db 15652 CTGGACCTGCTCTCTTTCACATACGAAAGATATGCGGTGGGATGGAACAGAAAG 15593  
 QY 739 oValGluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValI 759  
 Db 15592 -CTGGAGCTGCG-----AAAGACTTATGCTTACCAACCAATTTGCGACGATTCCTGAG 15539  
 QY 759 eArgGlyGlnIlePheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrG 779  
 Db 15538 AATGGGTGACAAAGGATACGCGCTTCGCTTCGATTCGGAACGGAGCGACGATGACGA 15479  
 QY 779 vAlaLysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSerGlu 799  
 Db 15478 GGGCTTCGTCGATGGAACCTTTCGAC-----CTTCTACCGACAA 15437  
 QY 799 eAsnAlaThrGlyHisGlnIleThrLeuLeuValProSerGluAspGlyAlaThrValLe 819  
 Db 15436 GCTATTGGCGGGGAGAGGCTTCGCTTCGCGAAGCCGCTGATACGGTTACGGTTC 15377  
 QY 819 uPheProIleArgProThrHisLeuGlyGluIleProIleThrValThrAlaLeuSerPr 839  
 Db 15376 GTTTCGCTCGATCTGTC----- 15358  
 QY 839 oThrAlaSerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIleGluLysSe 859  
 Db 15357 -TCCGATATGATGCTCTGCGGTGATGCTTGTATACCGAAGCAAGCATTCACGATGG 15299  
 QY 859 rTyrSerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysTr 879  
 Db 15298 AGAGCAACATCTGATGTCGACCTGCGCGCTACGAGCGAGCTGGAGACCATACCGTT 15239  
 QY 879 rLeuSerPheSerPheProProAsnThrValThrGlySerGluArgValGlnIleThrAl 899  
 Db 15238 GATTCTCTACGGCGAGACGCTCCAAACGGTG----- 15208  
 QY 899 aIleGlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTy 919  
 Db 15207 -----GATCTGATAGTCTA-----TTTCCCA 15185  
 QY 919 rGlyCysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLe 939  
 Db 15184 TCGGAGCGGAGACCGGCTATGGGTACATGACATTCACGCTG--GTGACAAATCCGCT 15128  
 QY 939 uThrLysLysLeuGln-----LeuThrArgAsnLeuLysGluLysAlaLeuSe 955  
 Db 15127 TTGGGTGCGCGTACAGGCTTCGCGCGATGATGATGATGATGATGATGATGATGATG 15068  
 QY 955 rPheMetArgGlnGlyTyrGlnArgLeuLeuLysTyrGlnArgGluAspGlySerPheSe 975  
 Db 15067 GGTGGCTTCGCGCTTCATTCGCAATGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 15009  
 QY 975 rAlaPheGlyAsnTyr-----AspProSerGlySerThrTripletLeuSerAlaPheVal-- 992  
 Db 15008 TGCATCTCTGCTTGGGTGAGAGATT--TCCGACATCTTGGCGGACGCTTCGAGATCA 14950  
 QY 993 -----LeuArgCysPhe 996  
 Db 14949 CTCGCTTACAAATATGCGCTTCGCTGCGAAGATTCGCTTGGCGCGGAGATGCTG 14890  
 QY 996 eLeuGluAlaAspProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTr 1016  
 Db 14889 GCC-GAAGCAACATCTGACGCTTTCAGGCTTCGCTGCTGCTGCTGCTGCTGCTGCTG 14831  
 QY 1016 r-----LeuLysGlyHisGlnLysSerAsnGlyLysPhe-- 1027



```
Db 14830 GACGAGCAGATGCTGATGACAACTGAAAAGCTGCAAAAGCTTGACGGTTCTGGGCC 14771
;
Qy 1028 -TTPASPPROGILYARGVALILEHISERGLUENGILYGLYASNLYSERPROVALIN 1047
;
Db 14770 TTGG-----CATCCTGAGATGTTCTTCCACGAT-----TA 14741
;
Qy 1047 rleuthralatylrilevalthserleu 1057
;
Db 14740 TCTGACGACTATGTGATGACGATGCTGGTA 14710
;
RESULT 14
US-09-311-352B-1
; Sequence 1, Application US/09311352B
; Patent No. 6329500
; GENERAL INFORMATION:
; APPLICANT: Webb, Donna J.
; APPLICANT: Gonias, Steven L.
; TITLE OF INVENTION: Transforming Growth Factor-beta Binding Site
; FILE REFERENCE: 00370-02
; CURRENT APPLICATION NUMBER: US/09/311,352B
; CURRENT FILING DATE: 1999-05-13
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 339
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-311-352B-1

Alignment Scores:
Pred. No.: 7,98e-16 Length: 339
Score: 234.00 Matches: 51
Percent Similarity: 58.56% Conservative: 14
Best Local Similarity: 45.95% Mismatches: 36
Query Match: 3.18% Indels: 10
Gaps: 3

US-10-020-095-4 (1-1428) x US-09-311-352B-1 (1-339)
Qy 660 TYRALAGLARGPhMeGcLUbnGluGluYHIsIleValAspIleHisAspPheSer 679
;
Db 28 TATGAGTACGATGATGATGGAAGAGCCATGCAAGCCCTGCTGATGTAAGAG----- 81
;
Qy 680 LeuGLYSerSerProHis-----ValArgLYNHisPheProGluThrTrpIleTrp 696
;
Db 82 -----CTTCACACGAGAGACCGGACCAAGTACTTCCCTGAGACATGATCTGG 129
;
Qy 697 LeuAspThrAsnMetGLYTrArgIleTYrGInGluPheGluValThrValProAspSer 716
;
Db 130 GATTGGTGCTGTGAAGTCAACAGAGGGGTGCTGAGGTAGAGTAAACAGTCCCTGACACC 189
;
Qy 717 IleThrSerTrpValAlaThrGLYpHeValIleSerGluAspLeuGLYleuGLYleuThr 736
;
Db 190 ATCACCAGATGAGAGGAGGAGGAGGCTTGTGCTGTGAGAGATGCTGAGACTTGTGTCTCT 249
;
Qy 737 ThrThrProValIGluGluGluAlaPheGlnPProPhePheIlePheLeuAsnLeuProTYr 756
;
Db 250 TCCACT---GCCCTCTCCGAGCCCTTCCAGCCCTTCTTTGTGAGCTCAATGCTTAC 306
;
Qy 757 SerValIleArgGLYGLYGLYGLYLeuPheAlaLeuGlu 767
;
Db 307 TCTGTGATTCGTGAGAGAGGCTTTCACACTCAG 339
;
RESULT 15
US-09-241-606-3
; Sequence 3, Application US/09241606
; Patent No. 6472140
; GENERAL INFORMATION:
; APPLICANT: Tanzi, Rudolph E.
; APPLICANT: Kovacs, Dora
; APPLICANT: Saunders, Aleister J.
; TITLE OF INVENTION: Alpha-2-Macroglobulin Therapies and Drug Screening Methods for
; TITLE OF INVENTION: Alzheimer's Disease
```

```
;
; FILE REFERENCE: 0609.4460003
; CURRENT APPLICATION NUMBER: US/09/241,606
; CURRENT FILING DATE: 1999-02-02
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 750
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(750)
; OTHER INFORMATION: A /LRP Binding Domain
US-09-241-606-3

Alignment Scores:
Pred. No.: 3.08e-13 Length: 750
Score: 215.50 Matches: 65
Percent Similarity: 45.08% Conservative: 45
Best Local Similarity: 26.64% Mismatches: 95
Query Match: 2.93% Indels: 39
Gaps: 7

US-10-020-095-4 (1-1428) x US-09-241-606-3 (1-750)
Qy 1164 IleMetArgTrpLeuSerArgGluNArgAsnSerLeuGLYpHeAlaSerThrGluAsp 1183
;
Db 28 ATCTGAGAGTGAATCAACAGACGACGAATGCCAGGCGGCTTCTCTCCACCCAGAGAC 87
;
Qy 1184 ThrThrValAlaLeuValAlaLeuSerGluPheAlaAlaLeuMetAsnThrGluArgThr 1203
;
Db 88 ACAGTGGTGGCTCTCCAGCTCTGTCCAAATATGAGCGGCCACATTTAC--AGACT 144
;
Qy 1204 AsnIleGlnValThrValIleThrGLYProSerSerProSerProLeuAlaValGln--- 1222
;
Db 145 GGAAGGCTGCACAGGTGACTATCCAGTCTTCAGGAGCATTTTCCAGCAAAATTCAGAGTG 204
;
Qy 1223 -----PrometAlaVal 1226
;
Db 205 GACAAACAATATGCGCTTACTGACAGAGCTTCTATTCGCCAGAGCTGCTGGGAATTC 264
;
Qy 1227 AsnIleSerAlaAsnGLYpHeGLYpHeAlaIleCYsGlnLeuAsnValValTYrAsnVal 1246
;
Db 265 AGCATGAAAGTACACAGGAGAGATGTGTACCTCCAGACCTCTTGAATACATATT 324
;
Qy 1247 LysAlaSerGLYSerSerArgArgArgSerIleGlnAsnGlnAlaPheAspLeu 1266
;
Db 325 -----CTCCAGAAAGAGAGATTCCTCTTT 351
;
Qy 1267 AspValAlaValLys-----GluAsnLysAspAspLeuAsn---HisValAspLeu 1282
;
Db 352 GCTTTAGAGATGCAACTCTGCTCAAACTTGATGTAAGAACCAAGCCACCAAGCTTC 411
;
Qy 1283 AsnValCYsThrSerPheSerGLYProGLY-----ArgSerGLYMetAlaLeuMet 1299
;
Db 412 CAATCTCCCTTAAGTGTGATGATACACAGGAGCGCGCTGCTCCCAATGCGCATCTGT 471
;
Qy 1300 GluValAsnLeuLeuSerGLYpHeMetValProSerGluAlaIleSerLeu----- 1316
;
Db 472 GATGTGAAGATGCTGCTGCTTCATTCCTCCAGCAACCAAGTGAATAATGCTTGAAGA 531
;
Qy 1317 SerGluThrValLysLysValAlaGluTYrAspHisGLYLysLeuAsnLeuTYrLeuAspSer 1336
;
Db 532 TCTAACCATGTGAGCGGACAGAACTGACGCAACCATGCTTGATTTACTTGAATAG 591
;
Qy 1337 ValAsnGluThrGlnPheCYsValAsnIleProAlaValArgAsnPheLysValSerAsn 1356
;
Db 592 GTGTCAATATCAGACCTAGCTGTCTTCAAGGTTCGCAAGATGCCAGTAAAGAT 651
;
Qy 1357 ThrGlnAspAlaSerValSerIleValAspTYrTYrGluProArgArgGlnAlaValArg 1376
;
Db 652 CTCAAAACGACCATGTAAGATGTATGATTACTAGAGAGGATGATGCTTGCATTCCT 711
```

Qy	1377	SeTyraSer	1380
Db	712	GAGTACATGCT	723

Search completed: January 17, 2004, 01:18:53  
Job time : 445 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: January 17, 2004, 01:19:00 ; Search time 982 Seconds  
(without alignments)  
5125.623 Million cell updates/sec

Title: US-10-020-095-4  
Perfect score: 7348  
Sequence: 1 MGGPPLTAAHLICVCTAL.....HSSVIFICFKLYEMELWL 1428

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 2324096 seqs, 1762381658 residues  
Total number of hits satisfying chosen parameters: 4648192

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO.epool\_p/US10020095/runat\_15012004\_163824\_18739/app\_query.fasta\_1.1607  
-DB=PubMed Applications NA -QMT=fastcap -SUFFIX=rmpb -MINMATCH=0.1  
-LOOPL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USRR=US10020095 @CGN\_1\_1.136 @runat\_15012004\_163824\_18739  
-NCPU=6 -ICPU=3 -NO MMAP -LARSQUERY -NEG SCORES=0 -WAIT -DSBIOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:  
1: /cgn2\_6/ptodata/2/pubna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/ptodata/2/pubna/PCT\_NEW\_PUB.seq:  
3: /cgn2\_6/ptodata/2/pubna/US06\_NEW\_PUB.seq:  
4: /cgn2\_6/ptodata/2/pubna/US06\_PUBCOMB.seq:  
5: /cgn2\_6/ptodata/2/pubna/US07\_NEW\_PUB.seq:  
6: /cgn2\_6/ptodata/2/pubna/PCTUS\_PUBCOMB.seq:  
7: /cgn2\_6/ptodata/2/pubna/US08\_NEW\_PUB.seq:  
8: /cgn2\_6/ptodata/2/pubna/US08\_PUBCOMB.seq:  
9: /cgn2\_6/ptodata/2/pubna/US09\_PUBCOMB.seq:  
10: /cgn2\_6/ptodata/2/pubna/US09B\_PUBCOMB.seq:  
11: /cgn2\_6/ptodata/2/pubna/US09C\_PUBCOMB.seq:  
12: /cgn2\_6/ptodata/2/pubna/US09\_NEW\_PUB.seq:  
13: /cgn2\_6/ptodata/2/pubna/US10\_PUBCOMB.seq:  
14: /cgn2\_6/ptodata/2/pubna/US10A\_PUBCOMB.seq:  
15: /cgn2\_6/ptodata/2/pubna/US10B\_PUBCOMB.seq:  
16: /cgn2\_6/ptodata/2/pubna/US10\_NEW\_PUB.seq:  
17: /cgn2\_6/ptodata/2/pubna/US60\_NEW\_PUB.seq:  
18: /cgn2\_6/ptodata/2/pubna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
------------	-------	-------------	-----------	----	-------------

1	4535.5	61.7	3033	10	US-09-833-381-1810	Sequence 1810, App
2	3360	45.7	2273	12	US-10-108-260A-953	Sequence 953, App
3	1441	19.6	4576	13	US-10-292-081A-7	Sequence 7, Appli
4	1441	19.6	4577	14	US-10-052-817-1	Sequence 1, Appli
5	1440	19.6	4422	9	US-09-873-403-4	Sequence 4, Appli
6	1440	19.6	4577	9	US-09-873-403-3	Sequence 3, Appli
7	1440	19.6	4577	10	US-09-880-107-2236	Sequence 2236, Ap
8	1440	19.6	4577	12	US-10-331-496A-19	Sequence 19, Appl
9	1440	19.6	4577	13	US-09-873-319-008	Sequence 408, App
10	1440	19.6	4577	13	US-09-960-706-654	Sequence 654, App
11	1440	19.6	4577	13	US-10-292-081A-3	Sequence 3, Appli
12	1440	19.6	4577	13	US-10-292-081A-5	Sequence 5, Appli
13	1440	19.6	4577	15	US-10-076-816-56	Sequence 56, Appl
14	1440	19.6	4809	13	US-10-240-965-178	Sequence 178, App
15	1438	19.6	4527	13	US-09-756-247-3	Sequence 3, Appli
16	1438	19.6	4527	13	US-09-756-247-5	Sequence 5, Appli
17	1429	19.4	4530	13	US-10-292-081A-2	Sequence 2, Appli
18	1422	19.4	4488	13	US-09-981-151A-9	Sequence 9, Appli
19	1418	19.3	4823	13	US-09-971-392-172	Sequence 172, App
20	1412.5	19.2	4595	10	US-09-917-800A-1531	Sequence 1531, Ap
21	1410	19.2	5092	13	US-10-006-285-1031	Sequence 404, App
22	1410	19.2	5092	13	US-09-971-429B-2	Sequence 2, Appli
23	1400	19.1	4615	10	US-09-880-107-3727	Sequence 3727, Ap
24	1390	18.9	4615	13	US-10-006-285-405	Sequence 405, App
25	1384	18.8	4677	13	US-10-316-253-266	Sequence 266, App
26	1384	18.8	4701	13	US-10-316-253-268	Sequence 268, App
27	1312	17.9	1300	13	US-10-133-013-223	Sequence 223, App
28	1274	17.3	4771	13	US-09-756-247-27	Sequence 27, Appl
29	1255	17.1	4426	9	US-09-925-301-552	Sequence 552, App
30	1157	15.7	4021	12	US-10-094-886-123	Sequence 123, App
31	1109.5	15.1	4501	12	US-10-094-886-125	Sequence 125, App
32	1072.5	14.6	5191	13	US-10-006-285-206	Sequence 206, App
33	991.5	13.5	5555	15	US-10-198-846-13979	Sequence 13979, A
34	988	13.4	5417	13	US-10-007-926A-413	Sequence 413, App
35	956	13.0	5087	15	US-10-001-076-111	Sequence 111, App
36	934	12.7	5067	11	US-09-964-824A-546	Sequence 546, App
37	934	12.7	5067	11	US-09-964-824B-18	Sequence 18, Appl
38	934	12.7	5067	13	US-10-101-510-316	Sequence 316, App
39	934	12.7	5067	15	US-10-001-076-3	Sequence 3, Appli
40	931	12.7	5067	9	US-09-875-519A-23	Sequence 23, Appl
41	930	12.7	875	13	US-10-160-162-48	Sequence 48, Appl
42	930	12.7	875	13	US-09-820-649-48	Sequence 48, Appl
43	930	12.7	5090	14	US-10-044-090-470	Sequence 470, App
44	925.5	12.6	6435	15	US-10-001-076-10	Sequence 10, Appl
45	921.5	12.5	4871	9	US-09-778-927A-18	Sequence 18, Appl

## ALIGNMENTS

RESULT 1  
US-09-833-381-1810  
; Sequence 1810, Application US/09833381  
; Patent No. US20020132090A1  
; GENERAL INFORMATION:  
; APPLICANT: Robison, Keith E.  
; TITLE OF INVENTION: No. US20020132090A1 Nucleic Acid and Protein Homologs  
; FILE REFERENCE: 5800-119  
; CURRENT FILING DATE: 2001-04-11  
; PRIOR APPLICATION NUMBER: 09/516,448  
; PRIOR FILING DATE: 2000-02-29  
; NUMBER OF SEQ ID NOS: 2050  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1810  
; LENGTH: 3033  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-833-381-1810

Alignment Scores:  
Pred. No.: 0  
Score: 4535.50  
Percent Similarity: 98.00%  
Length: 3033  
Matches: 880  
Conservative: 2

Best Local Similarity: 97.78% Mismatches: 1  
 Query Match: 61.72% Indels: 17  
 DB: 10 Gaps: 1  
 US-10-020-095-4 (1-1428) x US-09-833-381-1810 (1-3033)

QY 546 ValLeuYsIleProValGlnLeuValPheLysAsnYsIleYsLeuTyrTrpSerLys 565  
 DB 6 GTCCGAAAAATTCCTGTTCCAGCTGTTTTTAAAAATAAGATTAAGCTATATTGGAGTAAA 65  
 QY 566 ValLysAlaGluProSerGluLysValSerLeuArgIleSerValThrGlnProAspSer 585  
 DB 66 GTGAAAGCTGAAACCATCTGAGAAAGCTCTCTTAGGATCTGTGACACACCTGACCTCC 125  
 QY 586 IleValGlyIleValAlaValAspLysSerValAsnLeuMetAlaAspAsnAspIle 605  
 DB 126 ATAGTTGGGATTTGATGCTGTGACAAAGGTGAAATCTGATGAATGCTCTTAAGATATT 185  
 QY 606 ThrMetGluAsnValValHisGluLeuGluLeuTyrAsnThrGlyTyrTyrLeuGlyMet 625  
 DB 186 ACAATGGAAAAATGTGCTCATAGAGTTGGAACCTTTATACACAGGATATTATTAGGCATG 245  
 QY 626 PheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTyrPValLeuThrAspAlaAsn 645  
 DB 246 TTCAATGAATCTTTTGGCAGCTCTTTCAGGAATGTGACTCTGGATTAAGACAGATGCAAC 305  
 QY 646 LeuThrLysAspTyrTrpLeuAspGlyValTyrAspAsnAlaGluTyrAlaGluArgPheMet 665  
 DB 306 CTCACGAAAGATATTATGATGAGTGTATTATGACAAATGACAAATGCTGAGAGGTTATAG 365  
 QY 666 GluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeuGlySerSerProHis 685  
 DB 366 GAGGAAATGGAAGCATATTGTAGATATCATGACTTTCTTTGGGATGAGCATCCACAT 425  
 QY 686 ValArgLysHisPheProGluThrTrpIleTrpLeuAspThrAsnMetGlyTyrArgIle 705  
 DB 426 GTCCGAAAGCATTTTCCAGAGACTGTGATTTGGCTGACACCAACATGGGGTTAACAGATT 485  
 QY 706 TyrGlnGluPheGluValThrValProAspSerIleThrSerTyrPValAlaThrGlyPhe 725  
 DB 486 TACCAAGATTTGAAAGTACTGACTGATCTGATCTCACTCTTGGGATGCTACTGGTTTT 545  
 QY 726 ValIleSerGluAspLeuGlyLeuGlyLeuThrThrThrProValGluLeuGlnAlaPhe 745  
 DB 546 GTGATCTTGAGAGACTGGCTTGCTGACTTAACAATCACTCCAGTGGAGCTCCAGGCTTC 605  
 QY 746 GlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArgIlyGluGluPheAla 765  
 DB 606 CAACCATTTTTCATTTTGTGAACTTCCCTACTCTGTTATCAGAGGTGAAGAAATTGCT 665  
 QY 766 LeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrGluValLysValIleIleGlu 785  
 DB 666 TTGAAATTAATCTATATTCAATTAATTTGAAAGATGCACTGAGGTTAAGGTAATCATTTAG 725  
 QY 786 LysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsnAlaThrGlyHisGln 805  
 DB 726 AAAAGTGAATTTGATTTGATTTCTAATGACTTCAATGAATTAATGACACAGGCCACAG 785  
 QY 806 GlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPheProIleArgProThr 825  
 DB 786 CAGACCCCTTCTGTTCCAGTGGAGATGGGCAACTGTTCTTTTCCATCAGAGCCAAACA 845  
 QY 826 HisLeuGlyGluIleProIleThrValThrAlaLeuSerProThrAlaSerAspAlaVal 845  
 DB 846 CATCTGGAGAAATTCCTATACAGCTCACACTCTTTCACCCACTGCTTCTGATGCTATC 905  
 QY 846 ThrGlnMetIleLeuValIleValIleGluLysSerTyrSerGlnSerIleLeu 865  
 DB 906 ACCAGATGATTTTAAAGCTGACAGGATAGAAAATCATATTTCACATTCATCTTA 965  
 QY 866 LeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeuSerPheSerPhePro 885  
 DB 966 TTAGACTTGACTGACAAATAGGCTACAGAGTACCTTGAAAACTTGAGTTTCTCATTTCT 1025

QY 886 ProAsnThrValThrGlySerGluArgValGlnIleThrAlaIleGluAspValLeuGly 905  
 DB 1026 CCTTAATCACTGACTGGCAGCTGAAGAGTTCAATCACTGCAATTTGGAGATGTTCTTGGT 1085  
 QY 906 ProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGlyCysGlyGluGlnAsn 925  
 DB 1086 CTTTCCATCAATGGCTTAGGCTCATGATTCGATGCTTATGGCTGTGTAACGAAAC 1145  
 QY 926 MetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThrLysLysGlnLeu 945  
 DB 1146 ATGATTAATTTCTCCAAATATTACATTTGATTAATGACTTAAGAAACAACTG 1205  
 QY 946 ThrAspAsnLeuLysGluValAlaLeuSerPheMetArgGlnGlyTyrGlnArgGluLeu 965  
 DB 1206 ACAGATATTGAAAGAAAAAGCTCTTTCATTTATAGGCAAGGTTAACGAGAGAACTT 1265  
 QY 966 LeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyrAspProSerGlySer 985  
 DB 1266 CTCTATCAGAGGAAAGATGGCTCTTTCAGTGTCTTTGGGAATATGACCTTCTGGGAGC 1325  
 QY 986 ThrTrpLeuSerAlaPheValIleuArgCysPheLeuGluAlaAspProTyrIleAspIle 1005  
 DB 1326 ACTTGGTTCACCTTTTGTTTAAGATGTTCTTGAACCCGATCTTACATAGATATT 1385  
 QY 1006 AspGlnAsnValIleuHisArgThrTyrThrTrpLeuLysGlyHisGlnLysSerAsnGly 1025  
 DB 1386 GATCAGAAATGTTTACACAGAACATACACTGGCTTTAAAGACATCAGAAATCCAAACGT 1445  
 QY 1026 GluPheTrpAspProGlyArgValIleHisSerGluLeuGlnGlyLysAsnLysSerPro 1045  
 DB 1446 GAATTTGGGATCAGAGAAAGATGATTCATAGGACTTCAAGGTGGCAATTAAGATCCA 1505  
 QY 1046 ValThrLeuThrAlaTyrIleValThrSerLeuGlyTyrArgLysTyrGlnProAsn 1065  
 DB 1506 GTACACTTACACCTTATATTGTAACCTCTCTGGGATATGAAGATACAGCCTAAC 1565  
 QY 1066 IleAspValGlnGluSerIleHisPheLeuGluSerGluPheSerArgGlyIleSerAsp 1085  
 DB 1566 ATTGATGTGAAAGTATATCCATTTTGGATGCTGAATTCAGTAAAGGAATTTTCAGAC 1625  
 QY 1086 AsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGlySerProLysAlaLys 1105  
 DB 1626 AATTATCTCTAGCCCTTATTAATCTTATGATGTCTCATGAGTGGAGTCTTAAGCGAAG 1685  
 QY 1106 GluAlaLeuAsnMetLeuThrTrpArgAlaGluGlnGluGlyLysMetGlnPheTrpVal 1125  
 DB 1686 GAAGCTTTGAATGCTGACTTGGAGAGCAAGAACAAAGGTGCATGCAATTCGGGTG 1745  
 QY 1126 SerSerGluSerLysLeuSerAspSerTrpGlnProArgSerLeuAspIleGluValAla 1145  
 DB 1746 TCATCAGAGTCCAAACTTTTGACTCTCTGGACAGCAGCTCCCTGATATGAAAGTTGCA 1805  
 QY 1146 AlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGluGlyIleProIleMet 1165  
 DB 1806 GCGTATGCACTGCTCCACACTTCTTACAATTTCAAGCTTCAAGGGAATCCCAATTATAG 1865  
 QY 1166 ArgTrpLeuSerArgGlnArgAsnSerLeuGlyGlyPheAlaSerThrGlnAspThrThr 1185  
 DB 1866 AGGTGGCTAAGCAGGCAAAATAATACCTGGGGGTGTTTGCACTCTCAGGATACCACT 1925  
 QY 1186 ValAlaLeuLysValAlaLeuSerGluPheAlaAlaLeuMetAsnThrGluArgThrAsnIle 1205  
 DB 1926 GTGGCTTTAAGGCTGTGTCTGAATTTGCAAGCCTTAATGAATACAGAAAGACAAATATC 1985  
 QY 1206 GlnValThrValThrGlyProSerSerProSerPro----- 1217  
 DB 1986 CAAGTACCGTGACGGGGCTTACTCACAAGTCCGTGAAGTTTCGATTGACACACAC 2045  
 QY 1218 -----LeuAlaValAlaGlnProMetAlaValAlaAsnIle 1228  
 DB 2046 AACCGCTTACTCTTGACAGACAGAGACTTGTGTGTATCAGGCAATGGCAGTTAATATT 2105

```

QY 1229 SerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValValIleAsnValIleValAla 1248
DB 2106 TCGCAAAAGTGTGGATTTGCTATTGTCAGCTCAATGTTATATATGAAAGGCT 2165
QY 1249 SerGlySerSerArgArgArgSerIleGlnAsnGlnGluAlaPheAspLeuAspVal 1268
DB 2166 TCTGGATCTTCTAGAAAGAGAGATCTATCCAAATCAAGAACCTTTGATTTAGATGTT 2225
QY 1269 AlaValIleGlnAsnLysAspAspLeuAsnIleValAspLeuAsnValCysThrSerPhe 1288
DB 2226 GCGTAAAGAAATAAATGAATGATCTCAATCANTGATGATTTGATGATGTCACAGCTTT 2285
QY 1289 SerGlyProGlyArgSerGlyMetAlaLeuMetGluValAsnLeuLeuSerGlyPheMet 1308
DB 2286 TCGGAGCCCGGTAGAGAGTGCATGCTCTTATGAGATTAACTTATTAAGTGGCTTTAG 2345
QY 1309 ValProSerGluAlaIleSerLeuSerGluThrValIleValValGluValIleValIleGly 1328
DB 2346 GTGCTTCAGAAAGCAATTTCTCGAGCAGACAGTAAAGAAAGTGAATATGATCATGGA 2405
QY 1329 LysLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCysValAsnIleProAla 1348
DB 2406 AAACCTCAACTCATTTAGATTTCTGTAATGAAACCACTTTGTGTATATTTCTGCT 2465
QY 1349 ValArgAsnPhenylValSerAsnThrGlnAspAlaSerValSerIleValAspTyrTyr 1368
DB 2466 GTGAGAACTTTTAAAGTTTCAATATACCAAGATGCTTCAAGTGCATAGGATTAATCTAT 2525
QY 1369 GluProArgArgGlnAlaValAlaArgSerTyrAsnSerGluValIleLeuSerSerCysAsp 1388
DB 2526 GAGCCAAAGAGACAGCGGTGAGAAAGTTACACTGAAAGTGAAGTCTCTCTGTGAC 2585
QY 1389 LeuCysSerAspValGlnGlyCysArgProCysGluAspGlyValAspSerGlySerHis 1408
DB 2586 CTTTGAGAGATGTCAGGAGCTCCCTCTTGTGAGAGATGAGCTTCAAGCTTCCATCAT 2645
QY 1409 HisSerSerValIlePheIlePheCysPheValLeuLeuTyrPheMetGluLeuThrLeu 1428
DB 2646 CACTCTTCAGTCACTTTTATTTCTGTTCAAGCTTCTGACTTATGGAACCTTTGGCTG 2705

RESULT 2
US-10-108-260A-953
; Sequence 953, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1el full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 953
; LENGTH: 2273
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-108-260A-953

Alignment Scores:
Pred. No.: 0 Length: 2273
Score: 3360.00 Matches: 657
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 45.73% Indels: 0
DB: 12 Gaps: 0

US-10-020-095-4 (1-1428) x US-10-108-260A-953 (1-2273)
QY 1 MetGlnGlyProProLeuLeuThrAlaAlaHisLeuLeuCysValCysThrAlaAlaLeu 20
DB 66 ATGCAAGGCGCCACGCTCCGACCGCGCCACCTCTCTGCGTGTGACCGCGCGCTG 125
QY 21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGlyGly 40

```

```

DB 126 GCCGTGGCTCCCGGAGCTTCGTTTCTGGTGAACAGCCCAAGGATCATCAGGCCCGAGAGA 185
QY 41 AsnValIleThrIleGlyValGluLeuLeuGlnHisCysProSerGlnValThrValIleValAla 60
DB 186 AATGTGACTATTGGGGGTGAGCTTCTGGAAACCTGCCCCCTTCAACAGTGACTGTGAAGCG 245
QY 61 GluLeuLeuLysThrAlaSerAsnLeuThrValSerValLeuGluAlaGluGlyValPhe 80
DB 246 GAGCTGCTCAAGACAGCATCAACCTCACTGCTCTGCTCGAGAGCAAGAGAGATCTTT 305
QY 81 GluLysGlySerPhePheThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGlu 100
DB 306 GAAAAAGGCTCTTTAAAGACACTTACTCTTCATCATCACTCTGAACAGTGCAGATGAG 365
QY 101 IleTyrGluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThr 120
DB 366 AATTATGAGCTACGTTATACCGGAGTACCCAGATGAGATTTTATTTCTTAATGTATC 425
QY 121 ArgLeuSerPheGluThrTyrLeuArgIleSerValPheIleGlnThrAspLysAlaLeuTyr 140
DB 426 CGCTTATCATTTGAGACCAAGAGATATCTGCTTTCATTCAAACAGACAGAGCTTATAC 485
QY 141 LysProLysGlnGluValLysPheArgIleValThrLeuPheSerAspPheLysProTyr 160
DB 486 AAGCCAAAGCAAGAGTGAAGTTGCAATGTTTACACTTCTCAGATTTTAAAGCTTAC 545
QY 161 LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGlnThrLeu 180
DB 546 AAACCTCTTTTAAACATTTCTATTAAGAACCCCAATCAATTTGATCAACAGGTGTTG 605
QY 181 SerGlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIle 200
DB 606 TCACAAACAAAGATCTTGGAGTCACTTCCAAACTTTTCACTTATCTTCCATCATTA 665
QY 201 LeuGlyAspTyrPheIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGln 220
DB 666 CTTGTGATCGGTCTATTCAAGTTCAAGTGAATGACCAAGACATCATCAATCATTTGAG 725
QY 221 ValSerGluTyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSer 240
DB 726 GTTTTCAAGATATGATTTTACCAAAATTTGAAGTGAAGTTCGACAGACACATTAATGTTCT 785
QY 241 MetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrGlyLysProVal 260
DB 786 ATGAATCTTAAACATTTTAAATGTATCAACAGGCAAGATATCAATATGAGAGCCAGTG 845
QY 261 LysGlyAspValThrLeuThrPheLeuProLeuSerPheTyrPglyLysLysAsnIle 280
DB 846 AAAGAGAGCGTAAAGCTTACATTTTACCTTATCTTTGGGAAAGAAAGAAATATTT 905
QY 281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluLeuLys 300
DB 906 ACAAACAATTTAAATATGAATGATCTGCAAACTTCTTTTATATGATGAAGAGATGAAA 965
QY 301 AsnValMetAspSerSerAsnGlyLeuSerGluTyrIleLeuAspLeuSerSerProGlyPro 320
DB 966 AATGTAAATGATCTTCAAAATGAGACTTTCGAATACCTGGAATCATATCTCCCTGGAGACA 1025
QY 321 ValGluIleLeuThrThrValThrGluSerValThrGlyIleSerArgAsnValSerThr 340
DB 1026 GTAGAAATTTTAAACACAGTGAAGATCAAGTACAGGTATTTCAAGAAATGTAAGCACT 1085
QY 341 AsnValPhePheLysGlnHisAspTyrIleIleGluPhePheAspTyrThrValLeu 360
DB 1086 AATGTGTTCTTCAAGCAACATGATTAATCATCATTTGATTTTGAATATATCACTACTCTTGG 1145
QY 361 LysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsnGlnLeu 380
DB 1146 AAGCATCTCTCAACTTACAGCCACTGTGAAAGGTAATCTGATGAGGCAACCAACTG 1205
QY 381 ThrLeuGluGluArgArgAsnAsnValValIleThrValThrGlnArgAsnTyrThrGlu 400

```

```

Db      1206 ACTCTTGAGAAAGAGAAATTAATGTACTCATTAACAGTGCACAGAGAACTAATCTAGAG 1265
Qy      401 TyrTrpSerGlySerAsnSerGlyAsnGlnIlysmetGluAlaValGlnIlyLeuAsnTyr 420
Db      1266 TACTCGACCGGATCTTAACAGTGCAGAAATCAAGAAATGGAAGCTGTTCAGAAATAATTAAT 1325
Qy      421 ThrValProGlnSerGlyThrPheIlyValIleGlnPheProIleLeuGlnIlyAspSerSerGlu 440
Db      1326 ACTGCCCCCAAGTGAACCTTTTAAGATTGAATTCCTCAATCTCGAGAGATTCCAGTGGAG 1385
Qy      441 LeuGlnIlyValAlaTyrPheLeuGlnIlySerIlySerSerMetAlaValHisLeuLeuPhe 460
Db      1386 CTACAGTTGAAGCGCTTAATTCCTGTGTAGTAAAGTACAGCAAGGCACTCATAGTCTGTGTT 1445
Qy      461 LysSerProSerIlyThrTyrIleGlnIlyValThrAspGlnIlyAsnIleIlyValGly 480
Db      1446 AAGCTCTCTAGTAAGACATCAATCCAACTTAACAGATGAAATGAAATATTAAGGTGGGA 1505
Qy      481 SerProPheGlnIlyValIlySerGlyAsnIlyValIlySerGlnIlyLeuSerTyrMetVal 500
Db      1506 TCGCTTTTGAAGTGTGTAGTGTAGTGCACAAAGATTGAAGAGATTAGCTATATAGTGA 1565
Qy      501 ValSerArgIlyGlnIlyValAlaValGlnIlySerGlnIlyAsnSerThrMetPheSerLeuThr 520
Db      1566 GTATCCAGGGGACAGTTGTGTGTGTAGGAAACAAATTCACAAATGTTCTCTTTAACA 1625
Qy      521 ProGlnIlySerThrProIlyValIlyValIlyValIlyValIlyValIlyValIlyValIlyVal 540
Db      1626 CCAAGAAATTCCTGTGACTCCAAAGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1685
Qy      541 GluIleIleSerAspValIlyValIleProValGlnIlyValIlyPheIlyAsnIlyIleIly 560
Db      1686 GAAATTAATTAAGTATGTTCTTAATAATTCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1745
Qy      561 LeuTyrTrpSerIlyValIlyValIlyValIlyValIlyValIlyValIlyValIlyValIly 580
Db      1746 CTATATTGAGATTAAGTAAAGTGAAGCTGCAACATCTGTGAAGATCTCTGTGATCTCTGTG 1805
Qy      581 ThrGlnProAspSerIleValIlyIleValIlyValIlyValIlyValIlyValIlyValIlyVal 600
Db      1806 ACACAGCTGACTCATATGTTGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1865
Qy      601 AlaSerAsnAspIleThrMetGlnIlyValIlyValIlyValIlyValIlyValIlyValIlyVal 620
Db      1866 GCCCTATATGATTAATTAACATGGAATATGTCATGATGTCATGATGTCATGATGTCATGATG 1925
Qy      621 TyrTrpIlyGlnIlyMetPheMetAsnSerPheAlaValIlyPheGlnIlyCysGlyLeuTyrVal 640
Db      1926 TATTATTAGGCAATGTCATGATTAATTCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1985
Qy      641 LeuThrAspAlaAsnIlyValIlyValIlyValIlyValIlyValIlyValIlyValIlyVal 657
Db      1986 TTGACAGATGCAACCTCAGAGATTAATTAATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2036

RESULT 3
US-10-292-081A-7
; Sequence 7, Application US/10292081A
; Publication No. US20030162202A1
; GENERAL INFORMATION:
; APPLICANT: Kenneth David Becker
; APPLICANT: Gonul Velicelbi
; APPLICANT: Xin Wang
; APPLICANT: Randolph E. Tanzi
; APPLICANT: Lars Berttram
; APPLICANT: Aleister J. Saunders
; TITLE OF INVENTION: SINGLE NUCLEOTIDE POLYMORPHISMS AND MUTATIONS ON ALPHA-2-MACROGLOB
; FILE REFERENCE: 37481-3323
; CURRENT APPLICATION NUMBER: US/10/292,081A
; CURRENT FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: 60/337434
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0

```

```

; SEQ ID NO 7
; LENGTH: 4576
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-292-081A-7

Alignment Scores:
Pred. No.: 2,83e-133 Length: 4576
Score: 1441.00 Matches: 453
Percent Similarity: 46.20% Conserved: 264
Best Local Similarity: 29.19% Mismatches: 551
Query Match: 19.61% Indels: 284
DB: 13 Gaps: 60

US-10-020-095-4 (1-1428) X US-10-292-081A-7 (1-4576)
Qy      12 LeuLeuCysValCysThrAlaAlaLeuAlaValAlaProGlyProArgPheLeuValThr 31
Db      85 CTCTGTGCTCTCTGCCCCACAGACGCTCAGTCTGTGAAACCCGAGTATAGTGTCTG 144
Qy      32 AlaProGlyIleIleArgProGlyGlyAsnValThrIleGlyValGlnLeuLeuHis 51
Db      145 GTCCCTCTCTCTGCTCCACT---GAGACCACTGAGAGAGGCTGTCTCTTGAAGTAC 201
Qy      52 CysProSerGlnValThrValIlyValIlyValIlyValIlyValIlyValIlyValIlyVal 71
Db      202 CTGAATAGACAGTGAAGTGAAGTGTCTCTGAGTCTGTGAGGGAACAGAGAGCTC 261
Qy      72 --SerValIleGlnIlyValIlyValIlyValIlyValIlyValIlyValIlyValIlyVal 89
Db      262 TTCACTGACTGTGAGGGGAGGAGTGAACGTACTCCACTGTGTCCCTTC----- 309
Qy      90 LeuProSerLeuProLeuAsnSerAlaAspGlu-----IleTyrGlnLeuArgVal 106
Db      310 -----GCTGTCCCAAGTCTTCAATCAATGAGAGATATGTTCTCACTGTCCAAAGT 363
Qy      107 ThrGlyArgThrGlnIlyValIlyValIlyValIlyValIlyValIlyValIlyValIlyVal 126
Db      364 AAGGACCAACCCACAGAA-----TTTAAGAGGGGACCAACAGTATGTTAAGAAC 414
Qy      127 LysArgIleSerValPheIleGlnThrAspIlyAlaLeuTyrIlyProIlyGlnIlyVal 146
Db      415 GAGACAGCTGTGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 474
Qy      147 LysPheArgIleValIlyThrIlyPheSerAspPheIlyProTyrIlyThrIlySerIly 166
Db      475 AATTTGTGTGTGTCTCCATGATGAATACTTCAACCCCTGATGATGATGATGATGATGATGATG 534
Qy      167 Leu---IleLysAspProIlySerAsnLeuIleGlnIlyThrIlySerGlnIlySerAsp 185
Db      535 GTATACATTGAGATCCCAAGAAATCCGATCGCACAAATGGCAGATTCCAGTTAGAG 594
Qy      186 LeuGlyValIleSerIlyThrPheGlnIlySerSerHisProIleLeuGlnIlyAspTyrSer 205
Db      595 GGTGGCTCAAGCAATTTCTTTTCCCTCTCAAGAGCCCTTCCAGGAGCTCTCAAG 654
Qy      206 IleGlnValGlnValAsnAsp-----GlnThrTyrTyrGlnIlySerPheGlnIlySerGlu 223
Db      655 GTGGTGTGACAGAAATATCAAGTGAAGGACAGAGAC--CCTTACCGGTGAGGAA 711
Qy      224 TyrValLeuProIlyPheGlnIlyValIlyThrIlyGlnIlyThrProLeuTyrCysSerMetAsnSer 243
Db      712 TTTGTTCTTCCCAAGTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 771
Qy      244 LysHisLeuGlnIlyThrIleThrAlaIlyTyrThrIlyGlyIlyProValIlyGlyAsp 263
Db      772 GAAGAGATGAATATCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 831
Qy      264 ValThrIleThrPheLeuProLeu----- 271
Db      832 GTGACTGTGACATTTGCAGAAAGTATGATGACGCTTCGAGTCCAGGTTGAAGATTCA 891
Qy      272 -----SerPheTyrGlyLys 276

```

Db 892 CAGGCTTTCTGAGAAATTCAGTGCAGACAGCTAAACGCCAGTGCCTTCTATTCAGCA 951  
 QY 277 LysLeuAsnIleThrIleThrPheIleAsnGlySerAlaAsnPheSerPheAsnAp 296  
 Db 952 GTAAAT-----ACCAAGCTCTCCAGCTGAAG-----AAGAG 967  
 QY 297 GluGluMetIleAsnValMetAsnSerSerAsnGlyLeuSerGluIleuAsnLeuSer 316  
 Db 988 TATGAATGAAA-----CTTCACAGTGAAGGCCAGATCCAA 1023  
 QY 317 SerProGlyProValGluIleLeuThr-----ThrValThrGluSerVal 331  
 Db 1024 GAAGAAGAACACATGCTGTGAATGACTGGAAGCAGTCCAGTGAATTCACAAAGACATA 1083  
 QY 332 ThrGlyIleSerArgAsnValSerThrAsnValPhePheGluGlnHisAspTyrIleIle 351  
 Db 1084 ACCAAACTCTCA---TTGTGAAAGTGAGCTCACTTCGACAGGGA-----ATT 1131  
 QY 352 GluPhePheAspTyrThrThrValLeuLysProSerLeuAsnPheThrAlaThrValLys 371  
 Db 1132 CCCTTCTTGGG-----CAG 1146  
 QY 372 ValThrArgAlaAspGlyAsnGluLeuThrLeuGluGluArgArgAsnValValIle 391  
 Db 1147 GTGCGCTAGTATGATGGAAGGCGTCCCTATACCA-----ATAAAGTCATA 1194  
 QY 392 ThrValThrGluArgAsnTyrThrGluTyrTrpSerGlySerAsnSerGlyAsnGluLys 411  
 Db 1195 TTCATC---AGAGAAATGAGCAAACTATTAC---TCCATCTCCACGAGATGAG 1245  
 QY 412 MetGluAlaValGln---LysIleAsnTyrThr---ValProGlnSerGlyThrPhe 428  
 Db 1246 CATGCGCTGTATAGTCTCTATCAACACACCAATGATAGGTACTCTCTTACTGTT 1305  
 QY 429 LysIleGluPhePro-----IleLeuGluAsnSer 438  
 Db 1306 AGGGTCATATTACAGAGATGATCCCTGTACGCTACAGAGGTGTGCAGAAACAAC 1365  
 QY 439 SerGluLeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSer 458  
 Db 1366 GAAGAGGCACATCACTGCTAT----- 1389  
 QY 459 LeuPheLysSerProSerLysThrTyrIleGlnLeuLysThrArgAsnGluAsnIleLys 478  
 Db 1390 CTGTGTTCTCCCAAGCAAGAGCTTGTCCACTTGAGCCCATGTCTCATGAATACCC 1449  
 QY 479 ValGlySer-----ProPheGluLeuValValSerGlyAsnLys---Arg 492  
 Db 1450 TGTGGCCATATCTCAGACAGTCCAGGACATTATATCTGAATGAGGACCCGTGGGG 1509  
 QY 493 LeuLysGluLeuSer-----TyrMetValValSerArgGlyGlnLeuValAlaValGly 510  
 Db 1510 CTGAAGAGCTCTCTTATATATCTGAATATGGCAAGGAGCAATTGTCCGAACTGGG 1569  
 QY 511 -----LysGlnAsnSerThr-----MetPheSerLeuThr 520  
 Db 1570 ACTCATGACAGTCTGTGTGAAGCAGAGACATGAAGGCCATTTTCCATCTCAATCCCT 1629  
 QY 521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGluAsnAspGly 540  
 Db 1630 GTGAAGTCAAGACATCTCTCTCGCTCGCTGCTCATCATCATCTTTTACTACCGGG 1689  
 QY 541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560  
 Db 1690 GAGTATTTGGGATTCGCAAAATATGATTTGAATGTGTGCGCAACAAGGGTGAAT 1749  
 QY 561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580  
 Db 1750 TTGAGCTTCAAGCCCATCAACAAATCTCCAGCTCAAGCCCACTGGAAGTCAAGCG 1809  
 QY 581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMet--- 599

Db 1810 GCT---CTCAGTCCGCTGCGCCCTCCGCTGTGACCAAAAGCGTGCCTCATGAAG 1866  
 QY 600 -----AsnIleAsnSer----- 602  
 Db 1867 CCTGATGCTAGACTCTCGCGCTCTCGGTTTCAACTCTGCTACGAAAAAGAACTCACT 1926  
 QY 603 -----AsnAspIleThrMetGluAsnValValHisGluLeuGluLeu 616  
 Db 1927 GCGTTCCTCGGCGCTTTGAAATGACAGGACAAATGAAGACTGCATCATGTATATGTC 1986  
 QY 617 TyrAsnThrGlyTyr---TyrLeuGlyMetPheMetAsnSer-----PheAla 631  
 Db 1987 TATATTAATGGAATGCATATATCTCAGTATACAGTACAAAGAAAGATATGTATACGC 2046  
 QY 632 ValPheGlnGluCysGlyLeuThrValLeuThrAspAlaAsnLeuThrLysAspTyrIle 651  
 Db 2047 TTCCTAAGAGACATGGGCTTAAAGCAATTCACCAATCAAGATTCGTAAACCAAAATG 2106  
 QY 652 AspGlyValTyrAspAsnAlaGlu-----Tyr 660  
 Db 2107 TGTCCACAGCTTCAACAGATGAATGCATGAGTGAAGCTGACGTAGTTTAT 2166  
 QY 661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680  
 Db 2167 GAGTCAGATGTATATGGAAGAGCCCATGCACGCTGTGCATGTGAAGAG----- 2217  
 QY 681 GlySerSerProHis-----ValArgLysHisPheProGluThrTrpIleThrLeu 697  
 Db 2218 -----CCTCACAGGAGACCTTACGAAGATCTTCCGAGACATGATCTGGGAT 2268  
 QY 698 AspThrAsnMetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIle 717  
 Db 2269 TTGGTGTGTGTAACTCAGACAGTGTGCTGAGTAGAGTACATGCTCCCTACACATC 2328  
 QY 718 ThrSerTrpValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThr 737  
 Db 2329 ACCGATGTGAAGGCGAGGCGCTTGTGCTGTGAAGATGTGAGCTTGATCTCTTCC 2388  
 QY 738 ThrProValGluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSer 757  
 Db 2389 ACT---GCCTCTCTCGAGCTTCCAGGCTTCTTGTGTGAGCTCAACAATGCTTACTCT 2445  
 QY 758 ValIleArgGlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAla 777  
 Db 2446 GGTATTCGTGAGAGAGCTTACACTCAAGGCAAGGCTTAACTTACCTCCCAATGCC 2505  
 QY 778 ThrGluValLysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSer 797  
 Db 2506 ATCCGGGTCAAGTGCAGCTGGAAGCTCTCCGCTTCTTGAAGTGTCCAGTGGAGAAG 2565  
 QY 798 Glu-----IleAsnAlaThrGlyHisGlnGlnThrLeuLeuValProSer 812  
 Db 2566 GAACAAGGCGCTCACTGCATCTGTGCAAAAGCGGCGCA----- 2604  
 QY 813 GluAspGlyAlaThrValLeuPheProIleAspProThrHisLeuGlyGluIleProIle 832  
 Db 2605 -----ACTGTCTCTGGCAGTAACCCCAAGTATTTGGAATGTGAATTC 2652  
 QY 833 ThrValThrAlaLeuSer-----ProThrAla 841  
 Db 2653 ACTGTAGAGGCGAGAGCACTAGAGTCTCAGAGCTGTGTGGGACTGAGAGCTTCACT 2712  
 QY 842 -----SerAspAlaValThrGlnMetIleLeuValAlaGluGlyIle 856  
 Db 2713 CCTGAACAGGAAAGAAACACACATCAAGCTCTGTGTGTGAACCTGAAGAGACTA 2772  
 QY 857 GluLysSerTyrSerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThr 876  
 Db 2773 GAGAAAGAAACAACATTAATCTCTTGTCTCATATGAGTGTGAGAGTTCT----- 2826  
 QY 877 LeuLysThrLeuSerPheSerPheProAsnThrValIleThrGlySerGluArgValGln 896  
 Db 2827 ---GAAGATATATCCCTGAAGTGCACCAATATGTGTGAAGATCTGCCGAGCTTCT 2883



QY 897 ILeThrAlaIleGlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArg 916  
 Db 2884 GTCCACAGTTTGGGAGACATATTAGGCTCTGCATGCAGAAACACAAACAAATCTTCAG 2943  
 QY 917 MetProGlyGlyCysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeu 936  
 Db 2944 ATGGCTTATGGCTGGAGACAGAAATAGTCCCTTCTCTCAATCATATGATCTG 3003  
 QY 937 AspTyrLeuThrIleValLeuGlyLeuThrAspAsnLeuGlyValAlaLeuSerPhe 956  
 Db 3004 GATTATCTAAATGAACACAGAGCTTATCTCAGAGATGAACTCCAGGCGCATGGCTAT 3063  
 QY 957 MetArgGlnGlyTyrGlnArgGlyLeuLeuTyrGlnArgGlySerPheSerAla 976  
 Db 3064 CTCAAACAGCTGTTACAGAGACAGTGAATCAAAAGCTATAGTGGCTCTACAGCACC 3123  
 QY 977 PheGly---AsnTyrAspProSer---GlySerThrTyrLeuSerAlaPheValLeuArg 994  
 Db 3124 TTTGGGGAGCGCATATGGCAGAGAACCCAGGCAACCTGGCTCAGACCTTTGTTCTGAG 3183  
 QY 995 CysPheLeuGlnAlaAspProTyrIleAspIleAsnValLeuIleAsnGlyThrTyr 1014  
 Db 3184 ACTTTGGCCAGCTCGAGCTTACATCTTCATCGATGAGCAACATTAACCCAGCCCTC 3243  
 QY 1015 ThrTyrLeuValGlyIleGlnIleSerAsnGlyGluPheTyrAspProGlyArgValIle 1034  
 Db 3244 ATATGGCTCTCCAGAGCAGACAGACAAATGGCTGTTACAGAGCTCGTGCTCAGCTC 3303  
 QY 1035 HisSerGluLeuGlnGlyGlyAsnIleSerProValThrLeuThrAlaTyrIleValThr 1054  
 Db 3304 AACATGCGCATAAAGGGAGAGTAGATGAAGTGAAGCTCTCGGCTATATATACATC 3363  
 QY 1055 SerLeuLeuGlyTyrArgGlyTyrGlnProAsnIleAspValGlnIleSerIleHisPhe 1074  
 Db 3364 GCCCTTCTGAGANTCTCTCAGACTCAGCTGCTGTCGCAATGCCCTGTGTTTGC 3423  
 QY 1075 LeuGluSer-----GluPheSerArgGlyIleSerAspAsnTyrThr 1088  
 Db 3424 CTGAGATGAGCTGGAGACAGACAGACAGAGGAGCCATGGC---AGCAGTATATATAC 3480  
 QY 1089 LeuAlaLeuIleThrTyrAlaLeuSerSerValGlySer---ProValAlaLeuGlnAla 1107  
 Db 3481 AAACACACTGCTGGCTATGCTTTGGCTCGCAGAGTAAACAGCAAGAGAGAGAAATGA 3540  
 QY 1108 LeuAsnMetLeuThrTyrArgAlaGluGlnGlyMetGlnPheTyr----- 1124  
 Db 3541 CTCAGTCACTTAATGAGAAAGCTGTGAAAGAAACAACTGTTCATTGGAGCCCTC 3600  
 QY 1125 ValSerSerGluSerTyrLeuSerAspSerTyrGlnProArg-----SerLeuAspIle 1142  
 Db 3601 CAGAAACCAAGCAGCAGAGTGGGCGCATTTTACAGAACCCAGGCTCCTGCTGAGAGTG 3660  
 QY 1143 GluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGln-----ThrSer 1159  
 Db 3661 GAGATGACATCTTATGTCCTCTGCT---TATCTCAGGGCCAGCCAGCCCAACTCG 3717  
 QY 1160 GluGlyIlePro-----IleMetArgTyrLeuSerArgGlnArgAsnSerLeu 1175  
 Db 3718 GAGAGCCCTGAGCTTGCAGAACCAACATCTGTGAAGTGAACAGAGCAGCAAAATGCCAG 3777  
 QY 1176 GlyGlyPheAlaSerThrGlnAspThrThrValAlaLeuValAlaLeuSerGluPheAla 1195  
 Db 3778 GCGCGTTTCTCTCCACAGACAGACAGTGGGCTCTCCATGCTCTGTCGAAATATGGA 3837  
 QY 1196 AlaLeuMetAsnThrGlnArgThrAsnIleGlnValThrValThrGlyProSerSerPro 1215  
 Db 3838 GCAGCAGCATTTACC---AGAGCTGGGAAGCTGTGACAGCGATCATCATCTTCAAGG 3894  
 QY 1216 SerProLeuAlaValValGln----- 1222  
 Db 3895 ACATTTTCCAGCAATTCAGATGAGACAAACAAACCGCTGTATCTGCAGCAGGTCTCA 3954

QY 1223 -----ProMetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCys 1238  
 Db 3955 TTGCGAGCTGCTGGGGAATACAGCATGAAAGTACAGAGAGAGATGTGTACTCTC 4014  
 QY 1239 GlnLeuAsnValValTyrAsnValIleValSerGlySerSerArgArgAlaGlySerIle 1258  
 Db 4015 CAGACATCTCTGAAATACATATAT-----CTC 4041  
 QY 1259 GlnAsnGlnIleAlaPheAspLeuAspValAlaValIle-----GluAsnIleAsp 1275  
 Db 4042 CCAGAAAGAGAAAGATTCCTCCCTTGTAGAGTGCAGAGCTGCTGCTCAACTGTGAT 4101  
 QY 1276 AspLeuAsn---HisValAspLeuAsnValCysThrSerPheSerGlyProGly----- 1292  
 Db 4102 GAACCCAAAGCCACACAGCTTCCAAATCTCCCTTAAGTGTCAATTACACAGAGGCCGC 4161  
 QY 1293 ---ArgSerGlyMetAlaLeuMetGluValAsnLeuLeuSerGlyPheMetValProSer 1311  
 Db 4162 TCTGCTCCAAATGGGAGATGTGATGTAAGATGATGTCTGTGATTCCTCCCTGAG 4221  
 QY 1312 GluAlaIleSerLeu-----SerGluThrValIleValValGluTyrAspHisGly 1328  
 Db 4222 CCAGACGTAAATGCTTGAAAGATCTTAACCATGTAGCCGACAGAAAGTACAGCAGAAC 4281  
 QY 1329 IValLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCysValAlaIleProAla 1348  
 Db 4282 CATGTCTTATTAACCTTGATTAAGTGTCAATACAGACATGAGCTTGCTTACAGGTT 4341  
 QY 1349 ValArgAsnPheIleValSerAsnThrGlnAspAlaSerValSerIleValAspTyrTyr 1368  
 Db 4342 CTCAGAGATGCCAGTGAAGATCTGAAACGAGCATGATGAAGTCTATGATTAATCTAC 4401  
 QY 1369 GluProArgGlnAlaValAlaArgSerTyrAsnSer 1380  
 Db 4402 GAGACGATGATGATTGCAATTCCTGAGTACAAATGCT 4437

## RESULT 4

US-10-052-817-1  
 : Sequence 1, Application US/10052817  
 : Publication No. US20020114792A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Tanzi, Rudolph E.  
 : APPLICANT: Kovacs, Dora  
 : APPLICANT: Saunders, Aleister J.  
 : TITLE OR INVENTION: Alpha-2-Macroglobulin Therapies and Drug Screening Methods for  
 : FILE REFERENCE: 0609,4460005  
 : CURRENT APPLICATION NUMBER: US/10/052,817  
 : PRIOR FILING DATE: 2002-01-23  
 : PRIOR APPLICATION NUMBER: 09/241,606  
 : PRIOR FILING DATE: 1999-02-02  
 : PRIOR APPLICATION NUMBER: 09/148,503  
 : PRIOR FILING DATE: 1998-09-04  
 : PRIOR APPLICATION NUMBER: 60/093,297  
 : NUMBER OF SEQ ID NOS: 27  
 : SOFTWARE: PatentIn Ver. 2.0  
 : SEQ ID NO 1  
 : LENGTH: 4577  
 : TYPE: DNA  
 : ORGANISM: Homo sapiens  
 : FEATURE:  
 : NAME/KEY: sig\_peptide  
 : LOCATION: (44)..(112)  
 : NAME/KEY: CDS  
 : LOCATION: (44)..(4468)  
 : NAME/KEY: mat\_peptide  
 : LOCATION: (113)..(4468)  
 : US-10-052-817-1

Alignment Scores: 2,83e-133 Length: 4577  
 Pred. No.: 1441.00 Matches: 451  
 Score:

Percent Similarity: 46.208  
 Best Local Similarity: 29.064  
 Query Match: 19.614  
 DB: 14  
 Gaps: 59

US-10-020-095-4 (1-1428) x US-10-052-817-1 (1-4577)

QY 12 LeuLeuCyValCysThrAlaAlaLeuAlaValAlaProGlyProArgPheLeuValThr 31  
 |||||  
 |||||  
 DB 86 CTCTTGCTCTCTCCGACAGACGCGCTCAGTCTTGGAACCGCAGATATATGTTCTG 145

QY 32 AlaProGlyIleLeuLeuProGlyGlyAsnValThrIleGlyValGluLeuLeuGluHis 51  
 |||||  
 |||||  
 DB 146 GTCCCTCCCTCCGCTCCACACT---GAGACCACTGAGAGAGGCGTGTCTTCTGACCTCA 202

QY 52 CysProSerGlnValThrValIleValIleGluLeuLeuThrAlaSerAsnLeuThrVal 71  
 |||||  
 |||||  
 DB 203 CTGAATGAGACAGACTGACTGTAAGTCTTCTTGAGTCTGTCAAGGGAACAGAGACCTTC 262

QY 72 ---SerValLeuGluIleGluGly---ValPheGluIleGlySerPheLeuThr 89  
 |||||  
 |||||  
 DB 263 TTCACTGACCTGGAGCGGAGATGACGTATCTCACTGTGTGCGCTTC----- 310

QY 90 LeuProSerLeuProLeuAsnSerAlaAspGlu-----IleTyrGluLeuArgVal 106  
 |||||  
 |||||  
 DB 311 ---GCTGTCCCAAGTCTTCATCAATGAGAGATATGTTCTCTCACTGCTCAAGTG 364

QY 107 ThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThrArgLeuSerPheGluThr 126  
 |||||  
 |||||  
 DB 365 AAGAGCCAAACCCAAACA---TTTAAAGACCGGACCACTGATGTGTTAAGAC 415

QY 127 LysArgIleSerValPheIleGlnThrAspIleValAlaLeuTyrIleProLysGlnIleVal 146  
 |||||  
 |||||  
 DB 416 GAGGACAGCTGTCTTGTCCAGACAGCAAAATCAATCTACAAACAGGACAGACAGT 475

QY 147 LysPheArgIleValThrLeuPheSerAspPheLeuProTyrIleThrSerLeuAsnIle 166  
 |||||  
 |||||  
 DB 476 AATATTGTTGTTGTTCTCATGTAGTAAACTTTCACCCCTGATGATGATTCCACTA 535

QY 167 Leu---IleLysAspProLysSerAsnLeuIleGlnIleThrLeuSerGlnIleAsp 185  
 |||||  
 |||||  
 DB 536 GTATACATTCAAGATCCCAAGAGAAATCGCATGCGCAATGCGAGGTTTCCATTTAGG 595

QY 186 LeuGlyValIleSerIleThrPheGlnLeuSerSerIleProIleLeuGlyAspTyrSer 205  
 |||||  
 |||||  
 DB 596 GGTGGCTCAAGCAATTTCTTTCCCTCTCATTCAGAGCCCTTCAGAGGCTCTTCAAG 655

QY 206 IleGlnValGlnValAsnAsp-----GlnThrTyrTyrGlnSerPheGlnValSerGlu 223  
 |||||  
 |||||  
 DB 656 GTGGTGTACAGAAATCATAGGTGAGAGAGACAGACAC---CCTTTCACCGTGAGAGAA 712

QY 224 TyrValLeuProLysPheGlnIleValThrLeuGlnIleThrProLeuTyrCysSerMetLeuAsn 243  
 |||||  
 |||||  
 DB 713 TTGTGTTCTTCCAAAGTTTGAAGTACAAAGTACAGTCCCAAGATATATCACTTTGGA 772

QY 244 LysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProValIleGlyAsp 263  
 |||||  
 |||||  
 DB 773 GAAAGATATATGTATCACTGTGTGCTTATACATATGGAGAGCTGTCCCTGACAT 832

QY 264 ValThrLeuThrPheLeuProLeu----- 271  
 |||||  
 |||||  
 DB 833 GTGACGTGAGCAATTTGAGAAAGTATAGTACGCTTCGACCTGCCAGCGTGAAGTTCA 892

QY 272 --- 276  
 |||||  
 |||||  
 DB 893 CAGGCTTTCTGTGAGAAATTCAGTGCAGACCTTAACAGCCATGCTTCTTATACAGCA 952

QY 277 LysLysAsnIleThrIleThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAsp 296  
 |||||  
 |||||  
 DB 953 GTRAAA-----ACCAAGGCTTTCAGCTGAGAGAG-----AAGGAG 988

QY 297 GluIleMetLysAsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSer 316  
 |||||  
 |||||

DB 989 TATGAATGAAA-----CTTCACTAGAGCCAGATCCAA 1024

QY 317 SerProGlyProValGluIleLeuThr-----ThrValThrGluSerVal 331  
 |||||  
 |||||

DB 1025 GAAAGAGGACAGTGTGATTTGACTGGAAGCGCTCCAGTGAATATCAAGAACTCA 1084

QY 332 ThrGlyIleSerArgAsnValSerThrAsnValPhePheGlnHisAspTyrIleIle 351  
 |||||  
 |||||  
 DB 1085 ACCAAACTCTCA---TTTGTGAAGTGAAGCTCACTTTCGACAGAG-----ATT 1132

QY 352 GluPhePheAspTyrThrThrValLeuLysProSerLeuAsnPheThrAlaThrValLys 371  
 |||||  
 |||||  
 DB 1133 CCTTCTTTGGG-----CAG 1147

QY 372 ValThrArgAlaAspGlyAsnGlnLeuThrLeuGluGluIleArgArgAsnValValIle 391  
 |||||  
 |||||  
 DB 1148 GTGCCCTTAATGATGAGAAAGCGTCCCTATACCA-----ATAAATCATTA 1195

QY 392 ThrValThrGlnArgAsnTyrThrGluTyrTyrPheSerIleAsnSerGlyAsnGlnLys 411  
 |||||  
 |||||  
 DB 1196 TTCATC---AAGAGAAATGAAAGCAAACTATTAAC-----TCCAAATGTTACACGAGATGAG 1246

QY 412 MetGluAlaValGln---LysIleAsnTyrThr-----ValProGlnSerGlyThrPhe 428  
 |||||  
 |||||  
 DB 1247 CATGCCCTTGTACAGTCTCTATCAACACCAACGTTATGGGTACCTCTTACTGTT 1306

QY 429 LysIleGluPhePro-----IleLeuGluAspSer 438  
 |||||  
 |||||  
 DB 1307 AAGGTCAATTACAAAGATGATGATCCCTGTTACGGCTTACCACTGGGTGTCAGAAACAC 1366

QY 439 SerGluLeuGlnLeuLysAlaIleTyrPheLeuGlySerLysSerSerMetAlaValHisSer 458  
 |||||  
 |||||  
 DB 1367 GAAAGGACACATCAACACGCTTAT----- 1390

QY 459 LeuPheLysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLys 478  
 |||||  
 |||||  
 DB 1391 CTGTGTTCTCCCAAGCAAGAGCTTGTCCACCTTGAGCCCATGTCTCATGACCTACC 1450

QY 479 ValGlySer-----ProheGluLeuValIleSerGlyAsnLys-----Arg 492  
 |||||  
 |||||  
 DB 1451 TGTGGCATACTCAGACAGTCCAGGACATTATATATTCGAATGAGAGGACCCGTGGGG 1510

QY 493 LeuLysGluLeuSer-----TyrMetValIleSerArgIleGlnLeuValAlaValGly 510  
 |||||  
 |||||  
 DB 1511 CTGAAGAACTCTCTTTATATATCTGATATGCAAGAGGAGGATGTCCGAAGTGG 1570

QY 511 -----LysGlnAsnSerThr-----MetPheSerLeuThr 520  
 |||||  
 |||||  
 DB 1571 ACTCATGAGACTGTTGTGAGAGAGAGACATGAAAGGCCATTTTTCATCTCAATCCCT 1630

QY 521 ProGluAsnSerTyrThrProLysAlaCysValIleValTyrTyrIleGluAspAspGly 540  
 |||||  
 |||||  
 DB 1631 GTGAAGTACAGCAATGCTCTGTGCGTGGTGGTGCATCTATGCTGTTTACCTACCGGG 1690

QY 541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560  
 |||||  
 |||||  
 DB 1691 GACGTGATTTGGGATTTCTGCAAAATATATGATTTGAAAATTTGTCTGGCCACAAAGGTGAT 1750

QY 561 LeuTyrTyrSerLysValIleAlaGluProSerGluLysValSerLeuArgIleSerVal 580  
 |||||  
 |||||  
 DB 1751 TTGACCTTCAGGCCATCAACAAAGTCTCCAGCTCAGACGCCACCTCGAGTCAACAGG 1810

QY 581 ThrGlnProAspSerIleValIleValAlaValAspLysSerValAsnLeuMetAsn 600  
 |||||  
 |||||  
 DB 1811 GCT---CCTCAAGTCCGTGTGGCCCTCGTGTGGAGCAAAAGCGTCTGCTCATGTAGAG 1867

QY 601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeu---GluLeuTyrAsnThr 619  
 |||||  
 |||||  
 DB 1868 CCTGATGCTAGAGCTCTCGGCTCTCGGTTTACAACTGTCTACAGAAAGAGACCTCACT 1927

QY 620 GlyTyr-----TyrLeuGluLysMet 625  
 |||||  
 |||||  
 DB 1928 GGTCTTCCCTGGGCTTTGAATGACAGCAGATGAAGACTGCATCATGTATATGTTC 1987

Qy 626 PheMetAsnSer-----PheAla 631  
 Db 1988 TATATTATATGAAATACATATATCTCACTATCAAGTACAAATGAAAAGATATATGACAGC 2047  
 Qy 632 ValPheGlnGluCysGlyLeuThrValLeuThrAspAlaAsnLeuThrLysAspTyrIle 651  
 Db 2048 TTCCCTAGAGCATGTGGCTTTAAAGCATTCACCAATCCAAAGATTCGTAAACCCCAAAATG 2107  
 Qy 652 AspGlyValTyrAspAsnAlaGlu-----Tyr 660  
 Db 2108 TGTCCACAGCTTCAACAGTATGAAATGCAATGCACTGAGCTTACGTTAGCTTTTAT 2167  
 Qy 661 AlaGluArgPheMetGluGluAsnGluLysIleValAspIleHisAspPheSerLeu 680  
 Db 2168 GAGTCAGATGTATATGGAGAAAGGCGCATGACGCCCTGTCGATGTTGAAGAG----- 2218  
 Qy 681 GlySerSerProHis-----ValArgLysHisPheProGluThrTyrIleTyrLeu 697  
 Db 2219 -----CCTCACACGAGACCGCTACGAAAGTACTTCCCTGAGACATGATCTGGGAT 2269  
 Qy 698 AspThrAsnMetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIle 717  
 Db 2270 TTGGTGTGTAAACTACAGACGGGTGGCTGAGGTAGAGTAAACATCTCCAGACCCATC 2329  
 Qy 718 ThrSerTyrValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThr 737  
 Db 2330 ACCAGATGGAAAGCAGGGGCTTGTGCTGTGTAAGATGCTGAGACTTGTGATCTCTCC 2389  
 Qy 738 ThrProValGluLeuGluAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSer 757  
 Db 2390 ACT---GCCTCTCTCCGAGCCCTTCAGGCCCTTCTTGAGAGCTTACAACTCCTTACTCT 2446  
 Qy 758 ValIleArgGlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAla 777  
 Db 2447 GTGATTCGTGAGAGCCCTTCACTCACTCAAGGCGCTCAACTACTCTTCCCAATATGC 2506  
 Qy 778 ThrGluValLysValIleIleGlyLysSerAspLysPheAspIleLeuMetThrSerSer 797  
 Db 2507 ATCCGGGTCACTGTGTCAGCTGGAAGCCCTCTCCGCTTCTTGTGCTCCAGTGGAGAG 2566  
 Qy 798 Glu-----IleAsnAlaThrGlyHisGlnGlnThrLeuValProSer 812  
 Db 2567 GAAACAAGCGCTCACTGATCTGTGCAAAACGGCGGCA----- 2605  
 Qy 813 GluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGlyGluIleProIle 832  
 Db 2606 -----ACTGTCTCTGGGCAGTAAACCCAAAGTCAATTAGAAATGTAAATTTC 2653  
 Qy 833 ThrValThrAlaLeuSer-----ProThrAla 841  
 Db 2654 ACTGTGAGCGCAGAGCACTAGAGTCTCAAGAGCTGTGGAGCTGAGAGTCCCTTCACT 2713  
 Qy 842 -----SerAspAlaValThrGlnMetIleLeuValLysValGluGlyIle 856  
 Db 2714 CCTGAACAAGGAAAGAACACACAGTCACTGTTGTTGTAACCTGGAAGAGACTA 2773  
 Qy 857 GluLysSerTyrSerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThr 876  
 Db 2774 GAGAAAGAAACACATTCACCTCTACTTGTTCATCAGAGTGTGAGTTC----- 2827  
 Qy 877 LeuLysThrLeuSerPheSerPheProAsnThrValThrGlySerGlnArgValGln 896  
 Db 2878 ---GAAATATATTCCTGAAACTGCAACCAATGTGGTAGAAGATCTGCCGAGCTTCT 2884  
 Qy 897 IleThrAlaIleGlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArg 916  
 Db 2885 GTTCACAGTTTGGGAGACATATTAGCTCTGCAACGAAACACAAATCTTCTCAG 2944  
 Qy 917 MetProTyrGlyCysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeu 936  
 Db 2945 ATGCCCTATGCTGTGGAGACAGAAATATGTGCTCTTGTCTCAATCATATGATCTG 3004

Qy 937 AspTyrLeuThrLysLysGlnLeuThrAspAsnLeuLysGluValAlaLeuSerPhe 956  
 Db 3005 GATATCTTAATGAATAACAGCAGCTTACTCCAGAGATCAAGTCCAAAGCCATATGCTAT 3064  
 Qy 957 MetArgGlnIleTyrGlnArgGluLeuLeuTyrGlnArgGluLysArgLysPheSerAla 976  
 Db 3065 CTCAACACTGGTTACAGAGACAGTTGAATACAAACATATATATGCTCTCTCACAGACC 3124  
 Qy 977 PheGly---AsnTyrAspProSer---GlySerThrTyrPheSerAlaPheValLeuArg 994  
 Db 3125 TTGGGAGGAGATATGCGACGAAACAGGACACACCTGCTCACAGCTTGTTCGTAG 3184  
 Qy 995 CysPheLeuGluAlaAspProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyr 1014  
 Db 3185 ACTTTGCCAACCTGACGCTTACATCTTCATGATGAAACACACATATCCCAAGCCCTC 3244  
 Qy 1015 ThrTyrLeuLysGlyHisGlnLysSerAsnGlyGluPheThrAspProGlyArgValIle 1034  
 Db 3245 ATATGGCTCTCCAGAGCAGAAAGACAAATGCGTTTCAGAGCTTGAGTCACTGCTC 3304  
 Qy 1035 HisSerGluLeuGlnGlyAsnLysSerProValThrLeuThrAlaTyrIleValThr 1054  
 Db 3305 AACATGCCATTAAGGAGAGATAGAAAGATGAAGTACCTCTCCGCTATATCACCATC 3364  
 Qy 1055 SerLeuLeuGlyTyrArgLysTyrGlnProAsnIleAspValGlnGluSerIleHisPhe 1074  
 Db 3365 GCCCTCTGAGAGATCTCTCACAGTCACTCAACCTGTGTCCGAATGCCCTGTTTTC 3424  
 Qy 1075 LeuGluSer-----GluPheSerArgGlyIleSerAspAsnTyrThr 1088  
 Db 3425 CTGAGACACCTGGAAGACAGACAAAGAGGAGCAATGCG---AGCCATGTATATAC 3481  
 Qy 1089 LeuAlaLeuIleThrTyrAlaLeuSerSerValGlySer---ProLysAlaLysGluAla 1107  
 Db 3482 AAAGCATCTGTGGCTATGCTTTTGGCTGCGAGTACCAAGACAAAGAGAAAGTA 3541  
 Qy 1108 LeuAsnMetLeuThrTyrArgAlaGlnGluGlyGlyMetGlnPheTyr----- 1124  
 Db 3542 CTCAAGTCACTTATGAGAAAGCTGTGAAGAAAGACAACTGTGCATTTGGAGCGCCT 3601  
 Qy 1125 ValSerSerGluSerLysLeuSerAspSerTyrGlnProArg-----SerLeuAspIle 1142  
 Db 3602 CAGAAACCAAGGACAGCAGGGGCAATTTTACGAACCCAGGCTCCTCTGTGAGGTG 3661  
 Qy 1143 GluValAlaAlaTyrAlaLeuSerSerHisPheLeuGlnPheGln-----ThrSer 1159  
 Db 3662 GAGATGACATCTATGTGCTCTCGCT---TATCTCAGGCGCCAGCCAGCCCACTCG 3718  
 Qy 1160 GluGlyIlePro-----IleMetArgTyrLeuSerArgGlnArgAsnSerLeu 1175  
 Db 3719 GAGGACCTGACCTCTGCAACCAACATCTGTAAGTGAATCAGAAAGCAGAAATGCCAG 3778  
 Qy 1176 GlyGlyPheAlaSerThrGlnAspThrThrValAlaLeuLysValAlaLeuSerGluPheAla 1195  
 Db 3779 GGGCGTTTCTCTCCACCGAGACAGAGTGGTGTCTTCATCTCTGTCCAATATGGA 3838  
 Qy 1196 AlaLeuMetAsnThrGluArgThrAsnIleGlnValThrValThrGlyProSerSerPro 1215  
 Db 3839 GCGGCCACATTTAC---AGGACTGGAGAGGCTGCACAGGTACTATCCAGTCTTCAGGG 3895  
 Qy 1216 SerProLeuAlaValValGln----- 1222  
 Db 3896 ACAATTTCCAGAAATTCAGAGTGCACAAACATGCTGTATCTGACAGAGTCTCA 3955  
 Qy 1223 -----ProMetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCys 1238  
 Db 3956 TTGCGAGAGCTGCTGGAGATACAGCATGAAAGTGCACAGAGAGAGATGTGTACTCTC 4015  
 Qy 1239 GlnLeuAsnValValTyrAsnValLysAlaSerGlySerSerArgArgArgSerIle 1258  
 Db 4016 CAGACCTCTTGAATATCAATATT-----CTC 4042  
 Qy 1259 GlnAsnGlnIleAlaPheAspLeuAspValAlaValLys-----GluAsnLysAsp 1275

```

Db      4043 CCAAGAAAGAAAGTCCCTTGTAGAGTCAGACCTGCTCAAACTTGAT 4102
Qy      1276 Aspleuasn---HisValaspleuasnValCysThrSerPheSerGlyProGly----- 1292
Db      4103 GAACCCAAAGCCCAACGACGCTTCCAAATCTCCCTTAAGTGTCAATTACAGAGGAGCCG 4162
Qy      1293 ---ArgSerGlyMetAlaLeuMetGluValAsnLeuSerGlyPheMetValProser 1311
Db      4163 TCTGCTCCCAACATGGCGATGCTTAAGTGAAGATGCTCTGCTTCAATCCCGTAGAG 4222
Qy      1312 GluAlaIleSerLeu-----SerGluThrValIleValGluValGluValAspHisGly 1328
Db      4223 CCAACAGTAAATGCTTGAAGATTAACCATGTGAGCGGACAGAGTCAGACAGAAC 4282
Qy      1329 LysLeuAsnLeuIleLeuAspSerValAsnGluThrGlnPheCysValAsnIleProAla 1348
Db      4283 CATGCTTATTTACCTTGAATGAAGGTGATCAATCAGACACTGAGCTTGTTCACGGAT 4342
Qy      1349 ValArgAsnPhelyleValSerAsnThrGlnAspAlaSerValSerIleValAspTyr 1368
Db      4343 CTGCAAGATGTCCTCCAGTAAAGATCTCAACAGCAGCATAGTGAAGTCTATGATTA 4402
Qy      1369 GluProArgArgGlnAlaValArgSerTyrAsnSer 1380
Db      4403 GAGACGATGAGTTGCAATCGCTGAGTACAAATGCT 4438

```

## RESULT 5

```

US-09-873-403-4
/ Sequence 4, Application US/09873403
/ Patent No. US20020028207A1
/ GENERAL INFORMATION:
/ APPLICANT: SRISSAIVA, Pramod K
/ TITLE OF INVENTION: COMPLEXES OF ALPHA (2) MACROGLOBULIN AND ANTIGENIC
/ TITLE OF INVENTION: MOLECULES FOR IMMUNOTHERAPY
/ FILE REFERENCE: 8449-178
/ CURRENT APPLICATION NUMBER: US/09/873,403
/ CURRENT FILING DATE: 2001-06-04
/ PRIOR APPLICATION NUMBER: 09/625,139
/ PRIOR FILING DATE: 2000-07-25
/ PRIOR APPLICATION NUMBER: 60/209,266
/ PRIOR FILING DATE: 2000-06-02
/ SOFTWARE: FASTSEQ for Windows Version 3.0
/ SEQ ID NO 4
/ LENGTH: 4422
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-873-403-4

```

## Alignment Scores:

```

Pred. No.: 3,36e-133 Length: 4422
Score: 1440.00 Matches: 451
Percent Similarity: 46.20% Conservative: 266
Best Local Similarity: 29.06% Mismatches: 551
Query Match: 19,60% Indels: 284
DB: 9 Gaps: 59

```

US-10-020-095-4 (1-1428) x US-09-873-403-4 (1-4422)

```

Qy      12 LeuLeuCysValCysThrAlaIleValAlaValAlaProGlyProArgPheLeuValThr 31
Db      43 CTCCTTGCTCTCTGCGCCACAGACGCTCAGTCTCTGAAAACCGAGTATATGTTCTG 102
Qy      32 AlaProGlyIleIleArgProGlyGlyAsnValThrIleGlyValGluLeuLeuGlnHis 51
Db      103 GTCCCTCCCTGCTCCACACT---GAGACCACTGAAGAGGCTGTCTCTTGAGCTAC 159
Qy      52 CysProSerGlnValThrValIleValAlaGluLeuLeuValThrAlaSerAsnLeuThrVal 71
Db      160 CTGAATGAGACAGTGAAGTCTTGAAGTCTCTGAGCTGTCTCAGGGGAAACAGAGCCCTC 219
Qy      72 ---SerValLeuGluAlaGluGly---ValPheGluIleGlySerPheIleThrLeuThr 89

```

```

Db      220 TTCACTGACCTGAGAGCGAGATGACCTACTCCACTGTGTGCGCTTC----- 267
Qy      90 LeuProSerLeuProLeuAsnSerAlaAspGlu-----IleTyrGluLeuArgVal 106
Db      268 ---GCTGTCCCAAGCTTCAATGATGAGAGGTATGTTCTCTCACTGTCAAGG 321
Qy      107 ThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThrArgLeuSerPheGluThr 126
Db      322 AAAGACCAACCCAGAA---TTTAAAGAGCGGACCAAGTGAATGTTAAGAAC 372
Qy      127 LysArgIleSerValPheIleGlnThrAspIleValLeuTyrLysProLysGlnGluVal 146
Db      373 GAGCAGCTGCTGCTTGTCCAGACAGAAATCAATCAATCAAAACAGGAGCAGAGT 432
Qy      147 LysPheArgIleValThrLeuPheSerAspPheLysProTyrLysThrSerLeuAsnIle 166
Db      433 AAATTCGCTGTGTCCATGATGATAAATTTCCACCTGATGAGTGAATTCACCTTA 492
Qy      167 Leu---IleLysAspProLysSerAsnLeuIleGlnGlnThrPheSerGlnIleSerAsp 185
Db      493 GATACATTCAGATCCCAAGGAAATGCATCGACATGCGACAGATTTCAGTTAAG 552
Qy      186 LeuGlyValIleSerLysThrPheGluPheSerSerHisProIleLeuGlyAspTrpSer 205
Db      553 GGTGCGCTCAAGCAATTTCTTCTTCCCTCATAGAGCCCTTCAGAGGCTCCTAACAG 612
Qy      206 IleGlnValGlnValAsnAsp-----GlnThrTyrTyrGlnSerPheGlnValSerGlu 223
Db      613 GTGTGTGACAGAGAAATCAAGGTGAGAGCAGAGCAC---CCTTTCACCTCGAGAGA 669
Qy      224 TyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSerMetAsnSer 243
Db      670 TTTGTTCTTCCCAAGTTTGAAGTACAGTACAGTCCCAAGATATACCAATCTTGAA 729
Qy      244 LysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProValLysGlyAsp 263
Db      730 GAAAGAGAAATGATCATGAGTGTGCGCTATATACATATGGAAACGTCCTGTGACAT 789
Qy      264 ValThrLeuThrPheLeuProLeu----- 271
Db      790 GTGACTGTGAGCATTTGAGAAAGTATGATGACGCTTCGACGTCAGGAGTGAATTC 849
Qy      272 ---SerPheTrpGlyLys 276
Db      850 CAGGCTTTGTGAGAAATTCAGTGCAGACGCTAAACAGCCATGCTGCTTATACGAA 909
Qy      277 LysLysAsnIleThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAsp 296
Db      910 GTTAAA-----ACCAAGTCTTCCAGCTGAAGAG-----AAGAG 945
Qy      297 GluIleMetLysAsnValMetAspSerSerAsnGlyLysSerGlyTyrLeuAspLeuSer 316
Db      946 TATGAAATGAAA-----CTTACACATGAGGCGCCAGATCCAA 981
Qy      317 SerProGlyProValGluIleLeuThr-----ThrValThrGluSerVal 331
Db      982 GAAGAGGAAACAGTGTGAATGTAGCTGAAGGACGTCCAGTGAATATCAAGAACCA 1041
Qy      332 ThrGlyIleSerArgAsnValSerThrAsnValPhePheLysGlnHisAspTyrIleIle 351
Db      1042 ACCAAACCTCA---TTTGTGAAGTGAATGACATCTTCCACAGAGA-----ATT 1089
Qy      352 GluPhePheAspTyrThrThrValLeuLysProSerLeuAsnPheThrAlaThrValLys 371
Db      1090 CCTTCTTTGGG-----CAG 1104
Qy      372 ValThrArgAlaAspGlyAsnGlnLeuThrLeuGluGluArgAsnAsnValValIle 391
Db      1105 GTGCCCTAGTAGAGGAGGAGGCTCCCTATACCA-----AATTAAGTCAATA 1152
Qy      392 ThrValThrGlnArgAsnTyrThrGluTyrTrpSerGlySerAsnSerGlyAsnGlnLys 411

```



QY	1035	HisSerGIuLeuGIuNGIyAlaAsnLysSerProValThrLeuThrAlaTYrTlleValThr	1054
Db	3262	AAACATGCCATPAAAGGAGAGTAGTAAGAAATGAAGTACCCCTTCGCCCTAATATACCATC	3321
QY	1055	SerLeuLeuGIyTYrArgLysTYrGIuProAsnTlleAspValGIuInGIuSerTlleIAspHe	1074
Db	3322	GCCCTTCGGAGATTCCTCTCACAGTCACTACCCCTGTGTGTGCCAATGCCCTTTTGC	3381
QY	1075	LeuGIuSer-----GluPheSerArgGIyLiesAspAsnTYrThr	1088
Db	3382	CTGAGGTACAGCTGGAAAGACAGCACAGAAAGGGAGCATGCG---AGCAGATGATATATACC	3438
QY	1089	LeuAlaLeuLietmTYrAlaLeuSerSerValGIySer---ProLYaLysGIuAla	1107
Db	3439	AAAGCAGCTGGCGCTTATGTCTTTGGCCCTGGCAGGTATACAGACAAGAGAGAAAGAAATA	3498
QY	1108	LeuAsnMetLeuThrTYrArgAlaGIuInGIuGIuGIyGIyMetGIuPheTYrP-----	1124
Db	3499	CTCAAGTCACTTAATAGAGAAAGCTGTGAAGAAAGACAACCTGTTCATTGGAGAGCCCT	3558
QY	1125	ValSerSerGIuSerLysLeuSerAspSerTYrGIuProArg-----SerLeuAspIle	1142
Db	3559	CAGAAACCCAGAGCAGCGTGGGGGATTTTACGAAGCCAGCGCTCCCTGCTGAGAGTG	3618
QY	1143	GIuValAlaAlaTYrAlaLeuLeuSerHisPheLeuGIuPheGIu-----ThrSer	1159
Db	3619	GAGATGACATCTTATGTGCTCTCGCT---TATCTCAGCGCCAGCCAGCCCAACCTCG	3675
QY	1160	GIuGIyLysPro-----LleMetArgTYrLeuSerArgGIuArgAsnSerLeu	1175
Db	3676	GAGAGCTCGACCTCTGCACCAACATCTGTGAAGTGATCAGAAAGCAGCAAGAAATGCCAG	3735
QY	1176	GIyGIyPheAlaSerThrGIuAspThrThrValAlaLeuLysAlaLeuSerGIuLysPheAla	1195
Db	3736	GGCGGTTCCTCTCCACCCAGCAGACAGTGTGTGCTCCATGCTGTGTCCAAATATGGA	3795
QY	1196	AlaLeuMetAsnThrGIuArgTYrAsnLleGIuValThrValThrGIyProSerSerPro	1215
Db	3796	GCCCCCACTATTACC---AGGACTGGGAAGGCTGACAGGTACTATCCAGTCTTCAGGG	3852
QY	1216	SerProLeuAlaValAlaGIuIn-----	1222
Db	3853	ACAATTTCACGAAATTCCAATGGACACAAACAATGCCCTGTACTGACGACAGCTCTCA	3912
QY	1223	-----PrometAlaValaAsnLieserAlaAsnGIyPheGIyPheAlaLleCys	1238
Db	3913	TTGGCAGAGCTGCGCTGGAGATACAGCATGAAGAAATGACAGAGAAAGATGTGTACTCTC	3972
QY	1239	GIuLeuAsnValValTYrAsnValLysAlaSerGIySerSerArgArgArgSerTlle	1258
Db	3973	CAGACCTCTTGAATACAAATATT-----CTC	3999
QY	1259	GIuAsnGIuGIuAlaPheAspLeuAspValAlaValLys-----GluAsnLysAsp	1275
Db	4000	CCAGAAAGGAAGAGTTCCTCTTGGCTTTAGAGATGCGAGCTGCTGCTCAAACTTGTGAT	4059
QY	1276	AspLeuAsn---HisValAspLeuAsnValLysThrSerPheSerGIyProGIy-----	1292
Db	4060	GAACCCAAAGGCCACACACAGCTTCCAAATCTCCCTAAGTGCATTTACACAGAGAGCCGC	4119
QY	1293	---ArgSerGIyMetAlaLeuMetGIuValaAsnLeuLeuSerGIyPheMetValProSer	1311
Db	4120	TCTGCTTCCTCAACATGGCGATGTGATGTGAAGATGTGCTGTGCTTCAATCCCTGAAG	4179
QY	1312	GIuAlaLieserLeu-----SerGIuThrValLysLysValaGIuTYrAspHisGIy	1328
Db	4180	CCAAACAGGAAAAATGCTTGAAGAATCTAACCATGTGAGCCGAGACAGAAAGTCAGACGAC	4239
QY	1329	LysLeuAsnLeuTYrLeuAspSerValaAsnGIuThrGIuPheCysValaAsnLleProAla	1348
Db	4240	CATGCTTGATTTACCTTGATAGAGTGCACAAATGACACACTGAGACTTGTCTTCAACGGTT	4299
QY	1349	ValArgAsnPheLysValSerAsnThrGIuAspAlaSerValSerTlleValaAspTYrTYr	1368

[illegible]

```

Db      536 GTATACATTTCAGATCCCAAGAAATCGATCGCACATGCGACAGTTTCAGATTAGAG 595
      186 LeuGlyValIleSerLeuThrPheGlnLeuSerSerHisProIleLeuGlyAspTrpSer 205
      596 GGTGGCTTCAGAACATTTTCTTTCTTCCCTCATATAGAGCCCTTCCAGGGCTCCCTCAG 655
      206 IleGlnValGlnValAsnAsp-----GlnThrTyrTyrGlnSerPheGlnValSerGlu 223
      656 GTGGTGTACAGAAATAATCAGGTGGAGAGACAGAGCAGC-----CCTTACCGGTGAGGAA 712
      224 TyrValIleProLysPheGlnValThrLeuGlnThrProLeuTyrCysSerMetAsnSer 243
      713 TTTGTCTTCCCAAGTTTGAAGTACAGTAAACAGTCCCAAGATATACATCATCTTGAA 772
      244 LysHisLeuAsnGlyThrIleThrAlaLeuTyrThrTyrGlyValProValIleGlyAsp 263
      773 GAAGAGATGAATGATATAGTATGATGTGGCTTATACATATGGAGCCGTGCTCGACAT 832
      264 ValThrLeuThrPheLeuProLeu----- 271
      833 GTGACTGTGAGCATTTGCGAAGAAATAGTACGCTTCCGACTGCCAGCGTGAAGATTCA 892
      272 -----SerPheTrpGlyLys 276
      893 CAGGCTTTCTGTGAATAATTCAGTGCAGCTAAACAGCCATGCGCTCTCATACAGCA 952
      277 LysIleValAsnIleThrLeuThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAsp 296
      953 GTAAAA-----ACCAAGGTCTTCCAGCTGAAGAG-----AAGAG 988
      297 GluGluMetLysAsnValMetAspSerSerAsnGlyLeuSerGlyTyrLeuAspLeuSer 316
      989 TATGAATGAAA-----CTTCACACGTGAGCCCAAGATCAA 1024
      317 SerProGlyProValGlnIleLeuThr-----ThrValThrGlnSerVal 331
      1025 GAAGAGGAGAACAGTGGTGAATTGACTGAAGGAGCAGTCCAGTGAATATCAAGAAACCTATA 1084
      332 ThrGlyIleSerArgAsnValSerThrArgAsnValPhePheLysGlnHisAspTyrIleLe 351
      1085 ACCAAACCTCTCA-----TTGTGAAGTGCATCCACTTGCACAGGGA-----ATT 1132
      352 GluPhePheAspTyrThrThrValLeuLysProSerLeuAsnPheThrAlaThrValLys 371
      1133 CCCTTCTTGG-----CAG 1147
      372 ValThrArgAlaAspGlyAsnGlnLeuThrLeuGlnIleArgArgAsnValValIle 391
      1148 GTGGCGCTAGTAGATGGGAAAGCGTCCCTATACCA-----AATAAGTCTATA 1195
      392 ThrValThrGlnArgAsnTyrThrGluTyrTrpSerGlySerAsnSerGlyAsnGlnLys 411
      1196 TTCATC-----AAGAGAAATGAAGCAAACTATTAC-----TCCAAATGCTACCAAGGATAG 1246
      412 MetGluAlaValGln-----LysIleAsnTyrThr-----ValProGlnSerGlyThrPhe 428
      1247 CATGGCTTGTACACTTCTCATACACACACCAACGTTATGGTACTCTCTTACTTT 1306
      429 LysIleGluPhePro-----IleLeuGlnAspSer 438
      1307 AGGGTCATTTACAGAGATCGTAGTCCGTGTAACGGCTACAGTGGGTGTCAAGAAACAC 1366
      439 SerGluLeuGlnLeuLysAlaTyrPheLeuGlySerLysSerMetAlaValHisSer 458
      1367 GAAGAGGCACATCACACTGCTTAT----- 1390
      459 LeuPheLysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLys 478
      1391 CTGTGTTCTCCCAAGCAAGAGCTTGTCCACTTGGAGCCCATGTCTCAAGAACTACCC 1450
      479 ValGlySer-----ProPheGluLeuValIleSerGlyAsnLys-----Arg 492

```

```

Db      1451 TGTGGCATACTCAGACAGTCCAGGACCATTAATTTCTGAATGAGGACCCTGCTGGG 1510
      493 LeuLysGluLeuSer-----TyrMetValIleSerArgGlyGlnLeuValAlaValGly 510
      1511 CTGAAGAAGCTCTCTTTTATATCTATATATGCAATAGCAAGAGGACATTTGTCGAACGGG 1570
      511 -----LysGlnAsnSerThr-----MetPheSerLeuThr 520
      1571 ACTCATGAGACTGTTGTGAAGCAGAGAGACATGAAGGCCATTTTTCATCTCAATCCCT 1630
      521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGluAspArgGly 540
      1631 GTGAAGTCAACATTCCTCTGCTGCTGGTGTCTCATCTATGCTGTTTACTTACCCGG 1690
      541 GluIleIleSerAspValIleLysIleProValGlnLeuValPheLysAsnLysIleLys 560
      1691 GACGTGATTTGGGATTTCTGCAAAATATGATGTTGAAAATTTGTCGGCCAAACAGTGGAT 1750
      561 LeuTyrTrpSerLysValLysAlaGluProSerGlyLysValSerLeuArgIleSerVal 580
      1751 TTGAGCTTACAGCCCATCACAAGATCTCCAGCTCACAGCCCACTGCGAGTACAGCG 1810
      581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600
      1811 GCT-----CCTCAGTCTGCTGCGCCCTCGCTGTGTGACCAAAACGCTGCTTATAG 1867
      601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeu-----GluLeuTyrAsnThr 619
      1868 CTGTATCTGAGCTCTGCGGCTCTCGGTTTACAACTGCTTCCAGAAAAGAGACTCACT 1927
      620 GlyTyr-----TyrLeuGlyMet 625
      1928 GGTTCCTCGGCGCTTGAATGACACAGAGCATGAAGACTGCATCATCTCATATATGTC 1987
      626 PheMetAsnSer-----PheAla 631
      1988 TATATTTATGAAATCATCATATACTCCAGTATCAAGTACAATGAAGAAATATGTATGACG 2047
      632 ValPheGlnGluCysGlyLeuTrpValIleuThrAspAlaAsnLeuThrLysAspTyrIle 651
      2048 TTCTTAAGAGACATGGCTTTAAAGGACTTCAACCACTCAAAAGATTCTTAAACCAAAATG 2107
      652 AspGlyValTyrAspAsnAlaGlu-----Tyr 660
      2108 TGTCCACAGCTTCAACAGTATGAATAATCATGACCTGAAGGTCTACGTAGTATTAT 2167
      661 AlaGluArgPheMetGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680
      2168 GAGTCAGATGTAAATGGAGAGGCGCATGACGCTGTGTGATGTGAAG----- 2218
      681 GlySerSerProHis-----ValArgLysHisAspProGluThrTrpIleTrpLeu 697
      2219 -----CCTCACAGAGACCGTGAAGAGTCTCTCTGAGCATGATTTGGAT 2269
      698 AspThrAsnMetGlyTyrArgIleTyrGlnGluPheGlnValThrValProAspSerIle 717
      2270 TTGGTGTGTGAATACTAGACAGAGGGGTGCTGAGGTAGATTAACAGTCCCTGACACCATC 2239
      718 ThrSerTrpValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThr 737
      2230 ACCCAGTGAAGAGGAGGCGCTTCTGCTGTGAAATCTGACATCTGTGATCTTCTCC 2369
      738 ThrProValGluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSer 757
      2390 ACT-----GCTCTCTCCAGGCTTTCAGCCCTTCTTGTGAGAGCTTACAAAGCTTACTCT 2446
      758 ValIleArgGlyGlnGluPheAlaLeuGlnIleThrIlePheAsnTyrLeuLysAspAla 777
      2447 GTGATTTGTGAAGAGGCTTACACACTCAAGGACAGGCTCCAAACTCTTCCCAAAATGC 2506
      778 ThrGluValLysValIleIleGlyLysSerAspLysPheAspIleLeuMetThrSerSer 797
      2507 ATCCGGTCAGTGTGACGTGAAGGCTTCCGCTTCTGCTGTCCAGTGAAGAG 2566

```



QY 798 Glu-----11aenla1arhGlyHisGlnInhrLeuValProSer 812  
 DB 2567 GAACAAGCGCTCATGCTATCTGTGCAAGCGCGCA----- 2605  
 QY 813 GluAspGlyAlaIhrValLeuPheProIleArgProThrHisLeuGlyGluIleProIle 832  
 DB 2606 -----ACTGTCTCTGGGAGTAACCCCAAGTCATTAGGAATGGAATTTC 2653  
 QY 833 ThrValThrAlaLeuSer-----ProThrAla 841  
 DB 2654 ACTGTAGGCGCAGAGCCTAGAGTCTCAAGACTGTGGGACTGAGGTGCTTCAGTT 2713  
 QY 842 -----SerAspAlaValThrGlnMetIleLeuValIysAlaGluGlyIle 856  
 DB 2714 CTGGAACACGGAGAAAGACACAGTCATCAAGCTCTGTGGTTGACCTGAAGAGACTA 2773  
 QY 857 GluIysSerTyrSerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThr 876  
 DB 2774 GAGAGGAAACAAACATTCACTCCCTACTTGTCTCAGGTGAGGTTTCT----- 2827  
 QY 877 LeuIysThrLeuSerPheSerPheProProAsnThrValThrGlySerGluArgValGln 896  
 DB 2828 ---GAAGAATTATCCCTGAACCTGCCAACAATGTGTAGAGAAATCTGCCGAGCTTCT 2884  
 QY 897 IleThrAlaIleGlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArg 916  
 DB 2885 GTCTCAGTTTGGGAGACATATTAGGCTCTGCCATGCMAAACACACAAATCTTCTCCAG 2944  
 QY 917 MetProTyrGlyCysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeu 936  
 DB 2945 ATGCCCTATAGCTGTGGAGAGCAATATAGTCTCTTGTCTCTAACAATCTATAGTACG 3004  
 QY 937 AspTyrLeuThrIleIysIysGlnLeuThrAspAsnLeuLeuGluIysAlaLeuSerPhe 956  
 DB 3005 GATTATCTAATATAACACACAGCAGCTTACTCCAGAGCTCAAGTCCAGAGCATTTGCTAT 3064  
 QY 957 MetArgGlnGlyTyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAla 976  
 DB 3065 CTCACACGTGTATCCAGAGACAGTGTGAACATAACACATATATGCTCTACAGCACC 3124  
 QY 977 PheGly---AsnTyrAspProSer---GlySerThrTyrLeuSerAlaPheValLeuArg 994  
 DB 3125 TTGGGGAGAGATATGGCAGAGACAGGCAACCTGGCTCAGACCTTGTGTGAGG 3184  
 QY 995 CysPheLeuGlnAlaAspProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyr 1014  
 DB 3185 ACTTTGCCCAAGCTGAGCTTACATCTTATGATGAGACACATTAACCAAGCCCTC 3244  
 QY 1015 ThrTyrLeuIleGlyHisGlnIysSerAsnGlyGluPheTyrAspProGlyArgValIle 1034  
 DB 3245 ATATGCTCTCCAGAGGCGAAGAGACATGCTGTTTCAAGAGCTCTGGTATCAGCTC 3304  
 QY 1035 HisSerGluLeuGlnGlyIysAsnIysSerProValThrLeuThrAlaTyrIleValThr 1054  
 DB 3305 AACAAATGCAATAAAGGAGAGATGAAGATGACCTCTCCGCTATATACACATC 3364  
 QY 1055 SerLeuLeuGlyTyrArgIysTyrGlnProAsnIleAspValGlnGluSerIleHisPhe 1074  
 DB 3365 GCCCTTCTGAGATTCCTCTCAACATCACTCACTGTTGTCCGATGCGCTGTTTTC 3424  
 QY 1075 LeuGlnSer-----GluPheSerArgGlyIleSerAspAsnTyrThr 1088  
 DB 3425 CTGGAATCAAGCTCGAAGACACACACAAAGAGGAGCAATGCG---AGCATGTATATACC 3481  
 QY 1089 LeuAlaLeuIleThrTyrAlaLeuSerSerValGlySer---ProIysAlaIysGluAla 1107  
 DB 3482 AAGACACTGTGCTATGCTTTTGGCTCGGACAGTAAAGACAGACAAAGAGAAAGAA 3541  
 QY 1108 LeuAsnMetLeuThrTyrArgAlaGluGlnGluGlyMetGlnPheTyr----- 1124  
 DB 3542 CTCACATCACTTATATGAGAGAGCTGTGAAGAAAGAACAACTGTCTCATTTGGAGCGCCT 3601

QY 1125 ValSerSerGluSerIleLeuSerAspSerTyrGlnProArg-----SerLeuAspIle 1142  
 DB 3602 CAGAAACCCAGACGACCAATGGGGGATTTTTCAGAACCCCAAGCTCCCTCTGAGGTG 3661  
 QY 1143 GluValAlaIleTyrAlaLeuLeuSerHisPheLeuGlnPheGln-----ThrSer 1159  
 DB 3662 GAGATGACATCTTAATGTGTCTCTCGT---TATCTCACGGCCAGCCAGCACTCCG 3718  
 QY 1160 GluGlyIlePro-----IleMetArgTyrPheSerArgGlnArgAsnSerLeu 1175  
 DB 3719 GAGACCTACTCTTCGACCAACCAATCGTAAGTGAATCAGAGACAGAAATGCCAG 3778  
 QY 1176 GlyGlyPheAlaSerThrGlnAspThrThrValAlaLeuValAlaLeuSerGluPheAla 1195  
 DB 3779 GCGGTTTCTCTCCACCCAGAGACAGTGTGGTCTCTCAGTCTCTGTCCTCAATATGA 3838  
 QY 1196 AlaLeuMetAsnThrGluArgThrAsnIleGlnValThrValThrGlyProSerSerPro 1215  
 DB 3839 GCCGCCACATTATCC---AGGACTGGGAAGCTGCACAGTGTATCCAGTCTTCAGGG 3895  
 QY 1216 SerProLeuAlaValValGln----- 1222  
 DB 3896 ACAATTTCCAGCAAAATTCCAAGTGCACAAACAATGCGCTGTACTGCGAGAGTCTCA 3955  
 QY 1223 -----ProMetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCys 1238  
 DB 3956 TTGCCAGAGCTCTCTGGGAATACAGCATGAAGAGACAGAGAAAGATGTCTATCTTC 4015  
 QY 1239 GlnLeuAsnValValTyrAsnValIysAlaSerGlySerSerArgArgArgSerIle 1258  
 DB 4016 CAGACTCTCTGAAATACATATTT-----CTC 4042  
 QY 1259 GlnAsnGlnGluAlaPheAspLeuAspValAlaValIys-----GluAsnIysAsp 1275  
 DB 4043 CCAAGAAAGAAAGAGTCCCTTGTCTTAGAGTGCACACTGCTCTCAACTGTGTAT 4102  
 QY 1276 AspLeuAsn---HisValAspLeuAsnValCysThrSerPheSerGlyProGly----- 1292  
 DB 4103 GAACCCAAAGCCCAACAGCTTCCAAATCTCCATAGTGTACAGTTACAGAGGCCGC 4162  
 QY 1293 ---ArgSerGlyMetAlaLeuMetGluValAsnLeuLeuSerGlyPheMetValProSer 1311  
 DB 4163 TCTGCTCTCAACATGGCGCATCTGTATGTAAGATGTCTGTGCTTATCCCTCGAAG 4222  
 QY 1312 GluAlaIleSerLeu-----SerGluThrValIlyIysValGluTyrAspHisGly 1328  
 DB 4223 CCAACAGTAATAATCTTGAAGATCTTAACCATGTGAGCCGACAGAAAGTCAGCAAC 4282  
 QY 1329 LysLeuAsnLeuTyrIleAspSerValAsnGluThrGlnPheCysValAsnIleProAla 1348  
 DB 4283 CATGTCTTGATTTACTTGTATTAAGTGTCAATCAACACATGACATGAGCTTGTTCACGGTT 4342  
 QY 1349 ValArgAsnPheIysValSerAsnThrGlnAspAlaSerValSerIleValAspTyrTyr 1368  
 DB 4343 CTGCAAGATGTCCAGTAAAGATCTCAACACAGCAATGTAAGTGTATATTAACAC 4402  
 QY 1369 GluProArgArgGlnAlaValArgSerTyrAsnSer 1380  
 DB 4403 GAGACGATGAGTTTGCAATCGCTGATGATATGCT 4438  
 RESULT 7  
 US-09-880-107-2236  
 ; Sequence 2236; Application US/09880107  
 ; Patent No. US20020142981A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Horne, Darcil T.  
 ; APPLICANT: Vockley, Joseph G.  
 ; APPLICANT: Scherf, Iwe  
 ; APPLICANT: Gene Logic, Inc.  
 ; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer  
 ; FILE REFERENCE: 44921-5028-WO  
 ; CURRENT APPLICATION NUMBER: US/09/880,107  
 ; CURRENT FILING DATE: 2001-06-14

```

; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2236
; LENGTH: 4577
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 M11313
; US-09-880-107-2236

Alignment Scores:
Pred. No.: 3,566-133 Length: 4577
Score: 1440.00 Matches: 451
Percent Similarity: 46.20% Conservative: 266
Best Local Similarity: 29.06% Mismatches: 551
Query Match: 19.60% Indels: 284
DB: 10 Gaps: 59

US-10-020-095-4 (1-1428) x US-09-880-107-2236 (1-4577)
Qy 12 LeuLeuCyValCyThrAlaAlaValAlaProGlyProArgPheLeuValThr 31
Db CTCTTGCTGCTCTGCTCCAGAGAGCCCTCAGTCTCTGAAAACCGAGTATATGTTCTG 145
Qy 32 AlaProGlyIleIleArgProGlyGlyAsnValThrIleGlyValGluLeuGluHis 51
Db 146 GTCCCTCCCTCCCTCCACACT--GAGACCACTGAGAGGGCTGCTCTCTAGCTAC 202
Qy 52 CysProSerGlnValThrValLysAlaGluLeuLeuLysThrAlaSerAsnLeuThrVal 71
Db 203 CTGAATGACAGACGTACTGTAACTCTTCTTGAGCTGTCAAGGGAACAGAGCCTC 262
Qy 72 --SerValLeuGluAlaGluGly--ValPheGluLysGlySerPheLysThrLeuThr 89
Db 263 TTCACTGACCTGGAGGGCGAGATGACGTACTCCACTGTGCGCTTC----- 310
Qy 90 LeuProSerLeuProLeuAsnSerAlaAspGlu-----IleTyrGluLeuArgVal 106
Db 311 -----GCTGCTCCCAAGCTCTTCAATCAATGAGAGGTATGTTCTTCACTGTCCAGTG 364
Qy 107 ThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThrArgLeuSerPheGluThr 126
Db 365 AAAGACCAACCCCAAGAA-----TTTAAAGAGCGGACCAACAGTATGTTAAAGAC 415
Qy 127 LysArgIleSerValPheIleGlnThrAspLysAlaLeuTyrLysProLysGlnGluVal 146
Db 416 GAGGACACTGTGCTTGTCCAGACAGCAAAATCAATCTACAAACCAAGGCGACAGCTG 475
Qy 147 LysPheArgIleValThrLeuPheSerAspPheLysProTyrLysThrSerLeuAsnIle 166
Db 476 AAATTTCTGTGTTCTCATGATGATGAAACCTTCAACCCCTGAATGATGATTCCTCA 535
Qy 167 Leu---IleLysAspProLysSerAsnLeuIleGlnIleThrLeuSerGlnGlnSerAsp 185
Db 536 GTATATCAATTCAGATCCCAAGAAATGCAATCGCACATGCGACAGATTTCCAGTTAGAG 595
Qy 186 LeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIleLeuGlyAspTrpSer 205
Db 596 GGTGGCCCTCAAGCAATTTCTTCCCTCTCATAGAGCCCTTCCAGGCGCTCTACAG 655
Qy 206 IleGlnValGlnValAsnAsp-----GlnThrTyrTyrGlnSerPheGlnValSerGlu 223
Db 656 GTGGGTGACAGAGAAATCAGGTGAGAGCAGAGCAC--CCTTTCAACCGTGGAGGAA 712
Qy 224 TyrValLeuProLysPheGlnValThrLeuGlnThrProLeuTyrCysSerMetAsnSer 243
Db 713 TTGTGTTCTTCCCAAGTTGAAGTACAACTGACGCAAGAAAGATATCACTATCTGGA 772
Qy 244 LysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProValLysGlyAsp 263

```

```

Db 773 GAAGAGATGATATGATACAGTGTGCTTATACATATGGAAGGCTGTCCCTGACAT 832
Qy 264 ValThrLeuThrPheLeuProLeu----- 271
Db 833 GTGACTGTGACATTTGACAGAAAGTATAGACCTTCCAGCTCCACGGTGAAGATTCA 892
Qy 272 -----SerPheTrpGlyLys 276
Db 893 CAGGCTTCTGTGAGAAATTCAGTGGACAGCTAAACAGCATGCTGCTTCTATACAGCA 952
Qy 277 LysLysAsnIleThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAsp 296
Db 953 GTAATA-----ACCAAGTCTTCCAGCTGAGAGG-----AAGAG 988
Qy 297 GluGluMetLysAsnValMetLeuPheSerSerAsnGlyLeuSerGluTyrLeuAspLeuSer 316
Db 989 TATGAATGAAA-----CTTCACTAGAGCCCAAGATCCAA 1024
Qy 317 SerProGlyProValGluIleLeuThr-----ThrValThrGlnSerVal 331
Db 1025 GAAGAGGACACAGTGTGGAATTTGACTGGAAGGACGTCCAGTAATCAAGAACATA 1084
Qy 332 ThrGlyLysSerArgAsnValSerThrAsnValPhePheLysGlnHisAspTyrIleIle 351
Db 1085 ACCAACTCTCA--TTTGTGAAGTGAAGTCACTCATTTGACAGAGGA-----ATT 1132
Qy 352 GluPhePheAspTyrThrThrValLeuLysProSerLeuAsnThrAlaThrValLys 371
Db 1133 CCCTTCTTGGG-----CMG 1147
Qy 372 ValThrArgAlaAspGlyAsnGlnLeuThrLeuGluGluLysArgAsnValIle 391
Db 1148 GTGGCGCTACTAGATGGAGAAAGCGTCCATACCA-----AATAAGTCATA 1195
Qy 392 ThrValThrGlnArgAsnTyrThrGluTyrTrpSerGlySerAsnSerGlyAsnGlnLys 411
Db 1196 TTCATC--AGAGAAATGAGCAAACTATAC-----TCCATCTACCAAGCATGAG 1246
Qy 412 MetGluAlaValGln---LysIleAsnTyrThr-----ValProGlnSerGlyThrPhe 428
Db 1247 CATGGCTTATACAGTTCTCTATCAACACCAACAGTTATGGGTACTCTTCACTGTT 1306
Qy 429 LysIleGluPhePro-----IleLeuGluAspSer 438
Db 1307 AGGCTCAATTACAGAGATGTAGTCCCTGTACGGTACCAAGGGGTGACAGAAACAC 1366
Qy 439 SerGluLeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSer 458
Db 1367 GAAGAGCACATCACACTGCTAT----- 1390
Qy 459 LeuPheLysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLys 478
Db 1391 CTGTGTCTTCCCAAGCAAGACTTGTTCACCTTGAGCCCATGTCTCATGAACTACCC 1450
Qy 479 ValGlySer-----ProPheGluLeuValValSerGlyAsnLys-----Arg 492
Db 1451 TGTGGCATATCTCAGACAGTCCAGGACATTAATTTCTGAATGAGGACCCCTGTGGGG 1510
Qy 493 LeuLysGluLeuSer-----TyrMetValValSerArgGlyGlnLeuValAlaValGly 510
Db 1511 CTGAAGAGCTTCTCTTTATTTATCTGAATATGCAAGAGGAGGACTTGTCCAACTGGG 1570
Qy 511 -----LysGlnAsnSerThr-----MetPheSerLeuThr 520
Db 1571 ACTCATGAGCTGCTGTGAAGCAGGAAGACATGAGGGCCATTTTCCATCAATTCCT 1630
Qy 521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGluAspGly 540
Db 1631 GTGAAGTCAACATGTCTGTGCTGCGTGGTGTCTATCTATCTGTTTACTACCGGG 1690
Qy 541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560

```

Db 1691 GACGTCATTGGGATTCGCCAAATATGATGTTGAAAATTGCTGCGCAACAAGTGAT 1750  
 QY 561 LeuTYrTTPSerLySValIysAlaGluProSerGluLysValSerLeuArgIleSerVal 580  
 Db 1751 TTGAGCTTACGCCATCACAAGTCTCCAGCTTCACAGCCCACTGGAGATGACAGG 1810  
 QY 581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuAsn 600  
 Db 1811 GCT---CCTCAGTCCCTGCGCCCTCCGCTGCTGTGACCAAAAGCGTGTGCTCATGAAG 1867  
 QY 601 AlAserAsnAspIleThrMetGluAsnValValHisGluLeu---GluLeuTYrAsnThr 619  
 Db 1868 CCTGATGCTGAGCTCTGCGCGCTCTGCTTACAACTGCTACCAAAAAGAACCTCACT 1927  
 QY 620 GlyTYr-----TyrLeuGlyMet 625  
 Db 1928 GCGTTCCTGGGCGCTTGAATGACCAAGCATGAAGACTGCATCATGTCTAATATGTC 1987  
 QY 626 PheMetAsnSer-----PheAla 631  
 Db 1988 TATATTAAATGCAATCATACTACTCCAGTATCAAGTACAATGAAAAGATATGTACAC 2047  
 QY 632 ValPheGlnGluCysGlyLeuTyrPValLeuThrAspAlaAsnLeuThrIlysAspTYrIle 651  
 Db 2048 TTCTTAAGAGACATGGGCTTAAAGCATTACACCACTCAAAAGATTGTAAACCCAAATAG 2107  
 QY 652 AspGlyValTYrAspAsnAlaGlu-----Tyr 660  
 Db 2108 TGTCCACAGCTTCACACGATATGAATGATGACCTGAAGGCTCTACGTAGCTTTTAT 2167  
 QY 661 AlaGluArgPheMetGluLysAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680  
 Db 2168 GAGTCAGATGATATGGAGAGAGGCCATGCAAGCGCTGTGATGTGAAGAG----- 2218  
 QY 681 GlySerSerProHis-----ValArgIlyHisPheProGluThrTrpIleTrpLeu 697  
 Db 2219 -----CCTCACACGAGACCCGATCGAAAGTACTCTCCAGACATGAGTCTGGAGT 2269  
 QY 698 AspThrAsnMetGlyTYrArgIleTYrGlnGluPheGluValThrValProAspSerIle 717  
 Db 2270 TTGGTGTGTGTAACCTCAGACAGGGGCTGGCTGAGGTAGAGTAAACAGTCTCCAGACACATC 2329  
 QY 718 ThrSerTyrValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThr 737  
 Db 2330 ACCGAGTGAAGGACGAGGCGCTTCTGCGCTGTGAAGATGCTGGAATTTGATCTCTTC 2389  
 QY 738 ThrProValGluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTYrSer 757  
 Db 2390 ACT---GCCCTCTCCGAGCCTTCCAGCCCTTCTTGTGAGAGCTTTCATATGCTTACTCT 2446  
 QY 758 ValIleArgGlyGluGluPheAlaLeuGluIleThrIlePheAsnTYrLeuLysAspAla 777  
 Db 2447 GTGATTCGTGAAGAGCCTTTCACACTCAAGGCGCACGCTCTTAACCTTCCCAATGC 2506  
 QY 778 ThrGluValLysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSer 797  
 Db 2507 ATCCGGGTAGTGTGAGCTGGAAGGCTCTCCCGCTTCTTGCTGTCCCAAGTGAAGAG 2566  
 QY 798 Glu-----IleAsnAlaThrGlyHisGlnGlnThrLeuLeuValProSer 812  
 Db 2567 GAACAAGCCCTCACTGCTGTGTGCMAACGGGCGGCA----- 2605  
 QY 813 GluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGlyGluIleProIle 832  
 Db 2606 -----ACTGTGCTCGGGGAGTAAACCCCAAGTCAATTAGAATATGTGAATTC 2653  
 QY 833 ThrValThrAlaLeuSer-----ProThrAla 841  
 Db 2654 ACTGTAGAGGACAGGACCTAGAGTCTCAAGAGCTGTGGAGACTAGAGTGCCTTCAAGT 2713  
 QY 842 -----SerAspAlaValThrGlnMetIleLeuValIysAlaGluGlyIle 856  
 Db 2714 CCTGAACACGGAAGAAAGACACAGTCAATCAAGCTCTGTGTGTAACCTGAAGAGACTA 2773

QY 857 GluLysSerTYrSerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThr 876  
 Db 2774 GAGAAAGAAACAACATTCATCCCTACTTTCCTCATCAGGTGGAGTTCCT----- 2827  
 QY 877 LeuLysThrLeuSerPheSerPheProAsnThrValThrGlySerGluArgValGln 896  
 Db 2828 ---GAAAGATATATCCCTGAAACTGCCCAAAATGTGTAGAAGATCTGCCGAGCTTCT 2884  
 QY 897 IleThrAlaIleGlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArg 916  
 Db 2885 GTCTCAGTTTGGGAGACATATATAGCTCTGCAAGCAAAACACAAATTTCTCCAG 2944  
 QY 917 MetProTYrGlyCysGlyGluAsnMetIleAsnPheAlaProAsnIleTYrIleLeu 936  
 Db 2945 ATGCCCTTAGCTGTGGAGAGCAAGATATGTCTCTTGTCTCTTAACATCATATGACTG 3004  
 QY 937 AspTYrLeuThrLysLysLeuGluLeuThrAspAsnLeuLysGluLysAlaLeuSerPhe 956  
 Db 3005 GATATCTAAATGAAGAACAGAGCTTACTCCAGAGTCAAGTCCAGGCCATTGGCTAT 3064  
 QY 957 MetArgGlnGlyTYrGluArgGluLeuLeuTYrGlnArgGluAspGlySerPheSerAla 976  
 Db 3065 CTCACACTGTGTACCAAGACAGTGAACAAACATAATGATGAGTCTCTACAGCAC 3124  
 QY 977 PheGly---AsnTYrAspProSer---GlySerThrTrpLeuSerAlaPheValLeuArg 994  
 Db 3125 TTTGGGAGCGCATATGACAGAACCGAGGCAACCTGAGCTTACAGCTTGTGTGAGG 3184  
 QY 995 CysPheLeuGluAlaAspProTYrIleAspIleAspGlnAsnValLeuHisArgThrTYr 1014  
 Db 3185 ACTTTTGCCCAAGCTCGAGGCTTACATCTTCATGATGAAGACACATATCCCAAGCCCTC 3244  
 QY 1015 ThrTrpLeuLysGlyHisGlnLysSerAsnGlyGluPheTrpAspProGlyArgValIle 1034  
 Db 3245 ATATGGCTCTCCACAGGACAGAAAGCATGTCTGTTCAGAGAGCTTGGGTCACTGTGCTC 3304  
 QY 1035 HisSerGluLeuGlnGlyIysAsnLysSerProValThrLeuThrAlaTYrIleValThr 1054  
 Db 3305 AACATGCCATTAAGAGAGAGAGTAGAAGATGAGACCTCTCCGCTATATACCATC 3364  
 QY 1055 SerLeuLeuGlyTYrArgLysTYrGlnProAsnIleAspValGlnGluSerIleHisPhe 1074  
 Db 3365 GCCCTTCTGAGATTCCTCTCAGAGTCACTCAACCTGTGTGTCGGAATGCCCTGTTTGC 3424  
 QY 1075 LeuGlnSer-----GluPheSerArgGlyIleSerAspAsnTYrThr 1088  
 Db 3425 CTGAGTCAAGCCTGGAAAGACACACAAAGAGGAGCATGGC---AGCCATGTATATATAC 3481  
 QY 1089 LeuAlaLeuIleThrTYrAlaLeuSerSerValGlySer---ProLysAlaIysGluAla 1107  
 Db 3482 AAAGCATGTGGCTTAATGCTTTTGGCTTGGCAGGTAACCAAGACMAAGAAAGAAATGA 3541  
 QY 1108 LeuAsnMetLeuThrTrpArgAlaGluGlnGluGlyGlyMetGlnPheTrp----- 1124  
 Db 3542 CTCAGTCACTTAATAGAGAAAGCTGTGAAGAAAGCACTGTGTCCATTGGAGGCGCT 3601  
 QY 1125 ValSerSerGluSerLysLeuSerAspSerTrpGlnProArg-----SerLeuAspIle 1142  
 Db 3602 CAGAAACCAAGGACGACAGTGGGCAATTTTAAGAACCCAGAGCTCCCTGTGTAGAGTG 3661  
 QY 1143 GluValAlaAlaTYrAlaLeuLeuSerHisPheLeuGlnPheGln-----ThrSer 1159  
 Db 3662 GAGATGACATCTTAATGCTCTCTGCT---TATCTCAGGCGCCACGAGCCCAACCTTCG 3718  
 QY 1160 GluGlyIlePro-----IleMetArgTrpLeuSerArgGlnArgAsnSerLeu 1175  
 Db 3719 GAGAGCTGACCTCTGCAACCAACATCTGTGAAGTGAATCAAGAGCAGACAAATCCAG 3778  
 QY 1176 GlyIlyPheAlaSerThrGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAla 1195  
 Db 3779 GCGGCTTCTCTCTCCACCAAGACACAGTGTGTGCTCCATGTGCTCTGCCAAATATATGA 3838

```

Qy 1196 AlaLeuMetAnThrGlnArgThrAsnIleGlnValThrValThrGlyProSerSerPro 1215
Db 3893 GCCCGCACATTACC---AGGACTGGGAAGGCTGCACAGTACTATCTCAGTCTTCAAGG 3895
Qy 1216 SerProLeuAlaValValGln----- 1222
Db 3896 ACATTTCACGCAATTCGAAAGTGCACAAACATTCCTGTACTGCACAGGCTCA 3955
Qy 1223 -----PrometAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCyS 1238
Db 3956 TTGCCAGAGCTGCTCGGGGAATACGATCAAGTGAAGAGAGAGATGTCTACTCTC 4015
Qy 1239 GlnLeuAsnValValTyrAsnValIlySalSerGlySerSerArgArgArgSerIle 1258
Db 4016 CAGACTCTCTGAATCAATATTT-----CTC 4042
Qy 1259 GlnAsnGlnIlyAlaPheAspLeuAspValAlaValIlyS-----GlnAsnIlyAsp 1275
Db 4043 CCAGAAAGAGAAAGTCTCCCTTGTAGAGAGTGCAGCTGCTGCTCAAACTTGTAT 4102
Qy 1276 AspleuAsn---HisValAspleuAsnValCyThrSerPheSerGlyProGly----- 1292
Db 4103 GAACCCAAAGCCACACACAGCTCCAAATCTCCCTAAGTGTACATACAGGAGCCGC 4152
Qy 1293 ---ArgSerGlyMetAlaLeuMetGlnValAsnLeuSerGlyPheMetValProSer 1311
Db 4163 TCTGCTCTCAACATGCTGATGTGATGTAAGATGTCTGCTGCTTCAATCCCTGAG 4222
Qy 1312 GlnAlaIleSerLeu-----SerGlnThrValIlyIlyValGlnTyrAspHisGly 1338
Db 4223 CCACAGAGAAATGCTGAAAGATCTAACATGAGTGCAGCAGAGAGATGACAGCAAC 4282
Qy 1329 LysLeuAsnLeuTyrLeuAspSerValAsnGlnThrGlnPheCysValAsnIleProAla 1348
Db 4283 CATGCTGATTTTACCTTGATAGTGTCAATCAATCAACATGAGCTTCTTCAAGGTT 4342
Qy 1349 ValArgAsnPheIlyValSerAsnThrGlnAspAlaSerValSerIleValAspTyr 1368
Db 4343 CTGCAAGATGTCCCGTAAAGATCTCAACACGATGAGAAAGTCTATGATTTACTAC 4402
Qy 1369 GluProAlaArgGlnAlaValArgSerTyrAsnSer 1380
Db 4403 GAGACGATGAGTTGCAATCGTGAATCAATGCT 4438

RESULT 8
US-10-331-496A-19
; Sequence 19, Application US/10331496A
; Publication No. US20030228305A1
; GENERAL INFORMATION:
; APPLICANT: FRANTZ, GRETCHEN
; APPLICANT: HILLAN, KENNETH J.
; APPLICANT: PHILLIPS, HEIDI S.
; APPLICANT: POLAKIS, PAUL
; APPLICANT: SMITH, VICTORIA
; APPLICANT: SPENCER, SUSAN D.
; APPLICANT: WILLIAMS, P. MICKEY
; APPLICANT: WU, THOMAS D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF TUMOR
; FILE REFERENCE: P5014R1-PCT
; CURRENT APPLICATION NUMBER: US/10/331,496A
; CURRENT FILING DATE: 2002-12-30
; PRIOR APPLICATION NUMBER: US 60/345,444
; PRIOR FILING DATE: 2002-01-02
; PRIOR APPLICATION NUMBER: US 60/351,885
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: US 60/360,066
; PRIOR FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: US 60/362,004
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/366,869
; PRIOR FILING DATE: 2002-03-20

```

```

; PRIOR APPLICATION NUMBER: US 60/366,284
; PRIOR FILING DATE: 2002-03-21
; PRIOR APPLICATION NUMBER: US 60/368,679
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US 60/404,809
; PRIOR FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: US 60/405,645
; PRIOR FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 95
; SEQ ID NO 19
; LENGTH: 4577
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-331-496A-19

Alignment Scores:
Pred. No.: 3,56e-133 Length: 4577
Score: 1440.00 Matches: 451
Percent Similarity: 46.20% Conservative: 266
Best Local Similarity: 29.06% Mismatches: 551
Query Match: 19.60% Indels: 284
DB: 12 Gaps: 59

US-10-020-095-4 (1-1428) x US-10-331-496A-19 (1-4577)
Qy 12 LeuLeuCyValCysThrAlaAlaLeuAlaValAlaProGlyProArgPheLeuValThr 31
Db 86 CTCCTGCTCTCCTCGCCACAGACGCTCAGTCTCTGGAAGAACGCGATATGTTCTG 145
Qy 32 AlaProGlyIleIleArgProGlyGlyAsnValThrIleGlyValGlnLeuLeuHis 51
Db 146 GTCCTCTCCTCGCTCCACT---GAGACCATGAGAAAGGCTGTGCTCTGAGCTAC 202
Qy 52 CysProSerGlnValThrValIlyAlaGlnLeuLeuIlyThrAlaSerAsnLeuThrVal 71
Db 203 CTGAATGACAGAGTGTAGTGTCTCTGAGTCTGACGGGAAACAGAGGCTC 262
Qy 72 ---SerValLeuGlnIlyGly---ValPheGlnIlySerPheIlyThrLeuThr 89
Db 263 TTCACCTACTGAGGCGGAGATGACGTCTCCACTGCTCCCTC----- 310
Qy 90 LeuProSerLeuProLeuAsnSerAlaAspGlu-----IleTyrGlnLeuAlaVal 106
Db 311 -----GCTGCTCCAAAGTCTTCAATCAATGAGAGGTATGTTCTCACTGCTCAAGTG 364
Qy 107 ThrGlyArgThrGlnAspGlnIleLeuPheSerAsnSerThrArgLeuSerPheGlnThr 126
Db 365 AAAGACCAACCCAGAA-----TTTAAAGAGCGGACACAGATGATGTTAAGAAC 415
Qy 127 LysArgIleSerValPheIleGlnThrAspIlyAlaLeuTyrIlyProGlyAsnIlyVal 146
Db 416 GAGGACAGTGTGCTTGTTCAGACAGACAAATCAATCAACAAACAGGCGAGACGTG 475
Qy 147 LysPheArgIleValThrLeuPheSerAspPheIlyProTyrIlyThrSerLeuAsnIle 166
Db 476 AAATTGCTGTGCTCATGATGAAAACTTTCACCCCTGATGATGATTCACACTA 535
Qy 167 Leu---IleLysAspProIlySerAsnLeuIleGlnIlyThrPheLeuSerGlnIlyAsp 185
Db 536 GTATACCTTCAAGATCCCAAGAAATCGCATGACCAATGAGAGATTTCCAGTTAGAG 595
Qy 186 LeuGlyValIleSerIlyThrPheGlnLeuSerSerIlyProIleLeuGlyIlyAspTyrSer 205
Db 596 GGTGGCTCAAGCAATTTCTTTTCCCTTCAACAGACCTTCCAGGCTCTTCAAG 655
Qy 206 IleGlnValGlnValAsnAsp-----GlnThrTyrTyrGlnSerPheGlnValSerGln 223
Db 656 GTGGTGTAACAAGAAATCAGGTGAGAGACAGACAC---CCTTTCACCGTGAAGAA 712
Qy 224 TyrValLeuProIlyPheGlnValThrLeuGlnThrProLeuTyrCysSerMetAsnSer 243
Db 713 TTGTCTCTCCAAAGTTGAAGTAAAGTAAAGTAAAGTAAATCAATCAATCTTGGAA 772

```

Db	1691	GACGTGATTTGGGGATTCTGC	AAATATGATGTTGAAATATGTC	GGCCAAACAGGTGAT	1750	
Oy	561	LeuYrTrpSerLysVal	LyAlaGluProSerGluLysVal	SerLeuAlaGlyLeuSerVal	580	
Db	1751	TTGAGCTTCAGCCCATCA	CAAAAGCTCTCCAGCCTTC	ACGCCCAACGCTCGAGTCAACGC	1810	
Oy	581	ThrGlnProAspSerIle	ValGlyIleValAlaVal	AspLysSerValAsnLeuMetAsn	600	
Db	1811	GCT---	CCTCAGTCCGCTCGGCCCTCCG	TCGTGTGACCAAAAGCGTGTCTCATGAG	1867	
Oy	601	AlaSerAsnAspIle	ThrMetGluAsnValVal	HisGluLeu---GluLeuTyrAsnThr	619	
Db	1868	CCTGATGCTGAGCTCTCG	GGGCTCTCGGTTTACAACTG	CTACACAGAAAGAGACCTCACT	1927	
Oy	620	GlyTyr---	---	TyrLeuGlyMet	625	
Db	1928	GGCTTCCCTGGGCTTTGA	ATGACGAGACGATGAAAGCTG	CAATCGTCATTAATGTC	1987	
Oy	626	PheMetAsnSer---	---	PheAla	631	
Db	1988	TATATTAAATGGAATAC	ACATATACTCCAGTATCA	AGTCAATGAAAGATATGTCAGC	2047	
Oy	632	ValPheGlnLysCysGly	LeuTyrValLeuThrAspAla	AsnLeuThrLysAspTyrIle	651	
Db	2048	TTCTCTAAGGACATAG	GGCTTAAAGGCATTCACCACT	CAAAAGATTCTGTAACCAAAATG	2107	
Oy	652	AspLysValTyrAsp	AsnAlaGlu---	---	Tyr	660
Db	2108	TGTCCACAGCTTCAAC	AGTATGAAATGATCAGCTG	AGAGTCTAGGTTTAT	2167	
Oy	661	AlaGluLysrPheMetCys	GluLysGlnGlyHisIle	ValAspIleHisAspPheSerLeu	680	
Db	2168	GAGTCAGATGTAATAG	GGAAAGGCGCATGACCGCTG	ATGCAATGTGAAGAG---	2218	
Oy	681	GlySerSerProHis---	---	ValArgLysHisPheProGln	ThrTyrPileTyrLeu	697
Db	2219	-----CCTCACAGGAG	ACCGTACGAAGTACTTCC	TGACATGATCTGGGAT	2269	
Oy	698	AspThrAsnMetGlyTyr	ArgIleTyrGlnGluPheGluVal	ThrValProAspSerIle	717	
Db	2270	TTGTGTGTGTGAAC	CTCAGCAGGGGTGGCTGAGT	AGATGAACAGTCCCTGCACATC	2329	
Oy	718	ThrSerTrpValAla	ThrGlyPheValIleSerGlu	AspLeuGlyLeuGlyLeuThr	737	
Db	2330	ACCGAGTGAAGGACG	AGGGGCTTCTGCTGTGTA	MAATCTGACTGTGATCTCTCC	2389	
Oy	738	ThrProValGluLeu	GlnAlaPheGlnProPhePheIle	PheLeuAsnLeuProTyrSer	757	
Db	2390	ACT---GCCTCTCCG	AGCCTTCCAGCCTTCTT	TGTGAGGCTTCAATGCTTACTCT	2446	
Oy	758	ValIleArgGlyGlu	GluPheAlaLeuGlnIleThrIle	PheAsnTyrLeuLysAspAla	777	
Db	2447	GTCATGTGTGAGAG	GCCTTACACTCAAGCCG	CTCTTAATCACTTCCCAATGC	2506	
Oy	778	ThrGluValLysValIle	IleGluLysSerAspLysPheAspIle	LeuMetThrSerSer	797	
Db	2507	ATCCGGGTCAGTGTG	ACGTGGAAAGCCTCTCCG	CCCTTCTTGCTGTCCAGTGGAGAG	2566	
Oy	798	Glu---	---	IleAsnAlaThrGlyHisGln	GlnThrLeuLeuValProSer	812
Db	2567	GAAACAAGCGCTCAC	TGTCATCTGTGC	CAAAAGGGGGGCA-----	2605	
Oy	813	GluAspGlyAla	ThrValLeuPheProIleArgPro	ThrHisIleGlnGlyIleProIle	832	
Db	2606	-----ACTGTGTCT	GGGACGTAAACCCCAAGT	CATATTAGGAATGTGAATTTC	2653	
Oy	833	ThrValThrAlaLeu	Ser---	---	ProThrAla	841
Db	2654	ACTGTGAGCGCAGG	GCATAGACTCTCAAGAGCTGT	GTGGAGCTGAGGCTTCACTT	2713	
Oy	842	-----SerAspAla	ValThrGlnMetIleLeuVal	HisAlaGlnGlyIle	856	

Db 2714 CTGAACACGAGAGAAAGACACAGCTCATCAAGCTCTGTTGGTGAACCTGAAGACTA 2773  
 Qy GluYserSerThrSerGlnSerIleuLeuAspLeuThrAspAsnArgLeuGlnSerThr 876  
 Db 2774 GAGAAAGAAACAACTTCAACTCCCTACTTGTTCATCAAGGTGGAGGTTCT----- 2827  
 Qy LeuYSerThrLeuSerPheSerPheProProAsnThrValThrGlySerGluArgValGln 896  
 Db 2828 ---GAAAGATTATCCCTGAAACTGGCCACCAATGTGTAGAAGAAATGTGCCGAGCTTCT 2884  
 Qy IleThrAlaIleGlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArg 916  
 Db 2885 GCTCAAGTTTGGGACATATTAGCTCTGCCATGCAAAAACACACAAATCTTCTCCAG 2944  
 Qy MetProTyrGlyCysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeu 936  
 Db 2945 ATGCCCTATGGCTGTGGAGACAGAAATGTGCTCTTGTCTTCTTCAATCTATGTACTG 3004  
 Qy AspTyrLeuThrLysLysLysGlnLeuThrAspAsnLeuLysGluValAlaLeuSerPhe 956  
 Db 3005 GATTATCTAAATGAAACACAGAGCTTACTCCAGAGTCAAGTCCAAAGCCATTGGCTAT 3064  
 Qy MetArgGlnGlyTyrGlnArgGluLeuLeuTyrGlnArgGluArgGlySerPheSerAla 976  
 Db 3065 CTCAACACTGGTTACAGAGACAGTTGAACCTAACAAACCTATGAGCTCTCTACAGACC 3124  
 Qy PheGly---AsnTyrAspProSer---GlySerThrTyrLeuSerAlaPheValLeuArg 994  
 Db 3125 TTTGGGAGGCATATGACAGAAACAGGGCAACCTGGCTCAGAGCTTGTCTTCTGAG 3184  
 Qy CysPheLeuGlnAlaAspProTyrIleAspIleAsnGlnAsnValLeuIleArgThrTyr 1014  
 Db 3185 ACTTTGGCCAAAGCTCGAGCTACATCTTCATGCATGAAGACACATTAACCAAGCCCTC 3244  
 Qy ThrTyrLeuLysGlyHisGlnLysSerAsnGlyLeuPheTyrProGlyArgValIle 1034  
 Db 3245 ATATGGCTCTCCAGAGGCAAGAAAGCAATGTGCTTTCAGAGCTCTGGCTCACTGCTC 3304  
 Qy HisSerGluLeuGlnGlyValAsnLysSerProValThrLeuThrAlaTyrIleValThr 1054  
 Db 3305 AACAAATGCCATTAAGAGAGAGAGTAGAAGATGAAGTACCTCTCCGCTTATATCAATC 3364  
 Qy SerLeuLeuGlyTyrArgLysTyrGlnProAsnIleAsnValGlnGlnSerIleHisPhe 1074  
 Db 3365 GCCCTTGGAGATTCCTCTACAGTCACTCACTGCTGTGTCGGAATGCCCTGTTTGC 3424  
 Qy LeuGlnSer-----GluPheSerArgGlyIleSerAspAsnTyrThr 1088  
 Db 3425 CTGAGTACGCTCGAGAGACACACAAAGAGGAGCCATGCG---AGCCATGTATATACC 3481  
 Qy LeuAlaLeuIleThrTyrAlaLeuSerSerValGlySer---ProLysAlaLysGluAla 1107  
 Db 3482 AAAGCACTGCTGGCTTATGCTTTTGGCTTGGCAGGTATACAGACAAAGAGAGAGAGTA 3541  
 Qy LeuAsnMetLeuThrTyrArgAlaGlnGlnGlnGlyMetGlnPheTyr----- 1124  
 Db 3542 CTCAGTCACTTAATGAGAAAGAGCTGTAAAGAAACAATGTCATTTGGAGAGGCCCT 3601  
 Qy ValSerSerGluSerLysLeuSerAspSerTyrGlnProArg---SerLeuAspIle 1142  
 Db 3602 CAGAAACCCAGGACCCAGTGGGAGCTTTTACGAACCCCGAGGCTCCCTCTGAGAGGTG 3661  
 Qy GluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGln-----ThrSer 1159  
 Db 3662 GAGATGACATCTTATGCTCTCTCTGCT---TATCTCACGCGCCACGCCAGCCCAACTCG 3718  
 Qy GluArgIlePro-----IleMetArgTyrLeuSerArgGlnArgAsnSerLeu 1175  
 Db 3719 GAGGACCTGACCTCGCAACCAACATCTGTGAAGTGCATCAGAGACGACAAATGCCAG 3778  
 Qy GlyGlyPheAlaSerThrGlnAspThrThrValAlaLeuLysValAlaLeuSerGluPheAla 1195  
 Db 3779 GCGCGTTTCTCTCCACCCAGAGACAGTGTGTGCTCTCATGCTCTGTCCAATATATGA 3838

Qy 1196 AlaLeuMetAsnThrGluArgTyrAsnIleGlnValThrValThrGlyProSerSerPro 1215  
 Db 3839 GCGGCGCAATTTACC---AGGACTGGAGAGGCTGCACAGGTGACTATCCAGTCTTACGGG 3895  
 Qy 1216 SerProLeuAlaValValGln----- 1222  
 Db 3896 ACAATTTTCCGAAATTCGAAATGCGACACACAAATCGCTGTACTGACAGGCTCA 3955  
 Qy 1223 -----PrometAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCys 1238  
 Db 3956 TTGGCAGAGCTGCTGGGGAATACAGCATGAAGTGCAGAGAGAGATGTCTACTTC 4015  
 Qy 1239 GlnLeuAsnValValTyrAsnValLysAlaSerGlySerSerArgArgArgSerIle 1258  
 Db 4016 CAGACCTCTGGAATACATATTT-----CTC 4042  
 Qy 1259 GlnAsnGlnGluAlaPheAspLeuAspValAlaValLys-----GluAsnLysAsp 1275  
 Db 4043 CCAAGAAAGAGAGAGTTCCCTTGTGTGAGAGTGCAGACTGTGCTCAAACTTGAT 4102  
 Qy 1276 AspLeuAsn---HisValAspLeuAsnValCysThrSerPheSerGlyProGly----- 1292  
 Db 4103 GAACCCAAAGCCACACACAGCTTCCAAATCTCCCTTAAGTGTCACTTACACAGGAGCCG 4162  
 Qy 1293 ---ArgSerGlyMetAlaLeuMetGluValAsnLeuSerGlyPheMetValProSer 1311  
 Db 4163 TCGGCTCCCAACATGGGAGATCGTGAATGTGAAGATGTCTCTGGCTTCAATCCCGTGA 4222  
 Qy 1312 GluAlaIleSerLeu-----SerGluThrValLysLysValGluTyrAspHisGly 1328  
 Db 4223 CCAACAGTGAATAATGCTTGAAGATCTAACCAATGATGCGCGAGACAGACAGCACAG 4282  
 Qy 1329 LysLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCysValAsnIleProAla 1348  
 Db 4283 CATGCTTGAATTTACCTGTGAAGGTGTCAATCAAGACTGAGCTTGTCTTCAAGGTT 4342  
 Qy 1349 ValArgAsnPheLysLysSerAsnThrGlnAspAlaSerValSerIleValAspTyr 1368  
 Db 4343 CTCGAAGATCTCCAGTAAGAGATCTCAACACCAAGCCATGTGAAGTGTATGATTAATC 4402  
 Qy 1369 GluProArgArgGlnAlaValArgSerTyrAsnSer 1380  
 Db 4403 GAGACGATGATTTGCATTCGCTGAGTACAAATGCT 4438

RESULT 9  
 US-09-873-319-408  
 ; Sequence 408, Application US/09873319A  
 ; Publication No. US20030134324A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Mungier, William E.  
 ; APPLICANT: Kulkarni, Prakash  
 ; APPLICANT: Getzenberg, Robert H.  
 ; APPLICANT: Waga, Iwao  
 ; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic  
 ; TITL OF INVENTION: Hyperplasia Using Gene Expression Profiles  
 ; FILE REFERENCE: 44921-5029-US  
 ; CURRENT APPLICATION NUMBER: US/09/873,319A  
 ; EARLIER FILING DATE: 2001-06-05  
 ; EARLIER APPLICATION NUMBER: US 60/223,323  
 ; NUMBER OF SEQ ID NOS: 755  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 408  
 ; LENGTH: 4577  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; OTHER INFORMATION: Genbank Accession No. US20030134324A1 M11313  
 US-09-873-319-408  
 Alignment Scores:

Pred. No.: 3,566-133 Length: 4577  
 Score: 1440.00 Matches: 451  
 Percent Similarity: 46.20% Conservative: 266  
 Best Local Similarity: 29.06% Mismatches: 551  
 Query Match: 19,604 Indels: 284  
 DB: 13 Gaps: 59

US-10-020-095-4 (1-1428) x US-09-873-319-408 (1-4577)

QY 12 LeuLeuCyValCyThrAlaLeuAlaValAlaProGlyProArgPheLeuValThr 31  
 |||||  
 |||||  
 Db CTCTTGTCCTCTCTGCGCCACAGAGCGCTCATGCTCTGGAACCGAGATATGTTCTG 145  
 QY 32 AlaProGlyIleLeuArgProGlyGlyAsnValThrIleGlyValGluLeuLeuGlnHis 51  
 |||||  
 |||||  
 Db GTCCCTCCCTGCTCCACACT--GAGACCACTGAGAGAGGCGTGTCTCTGAGACTAC 202  
 QY 52 CysProSerGlnValThrValValAlaGluLeuLeuThrAlaSerAsnLeuThrVal 71  
 |||||  
 |||||  
 Db CTGAATGAGACAGTGAAGTGAAGTCTCTGAGTCTGTCAAGGGAACAGAGACCTTC 262  
 QY 72 ---SerValLeuGluAlaGluGly---ValPheGluLeuGlySerPheLeuThr 89  
 |||||  
 |||||  
 Db TTCACTGACTGAGCGAGGAGAGATGACGTACTCCACTGTGCGCTTC----- 310  
 QY 90 LeuProSerLeuProLeuAsnSerAlaAspGlu-----IleTyrGluLeuArgVal 106  
 |||||  
 |||||  
 Db GCTGTCCCAAGTCTTCATTCATTCAGAGAGTATGTTCTCACTGTGCAAGTG 364  
 QY 107 ThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThrArgLeuSerPheGluThr 126  
 |||||  
 |||||  
 Db AAAGAGCAACCCCAAGAA-----TTTAAAGACCGAGCAAGTATGTTAGTTAAGAC 415  
 QY 127 LysArgIleSerValPheIleGlnThrAspLysAlaLeuTyrIleProGlyGlnGluVal 146  
 |||||  
 |||||  
 Db GAGAGACGTGCTGCTTGTCCAGACAGCAATCATCTACAAACAGGAGAGACAGT 475  
 QY 147 LysPheArgIleValThrLeuPheSerAspPheLysProTyrIleThrSerLeuAsnIle 166  
 |||||  
 |||||  
 Db AAATTTCCGTGTGTCTCCATGATGAATACTTCACTCCCTGATGATGATGATTCACAT 535  
 QY 167 Leu---IleLysAspProLysSerAsnLeuIleGlnIleTyrLeuSerGlnIleAsp 185  
 |||||  
 |||||  
 Db GATATCATTCAGAGATCCCAAGAGAAATCGCATCGACATGGAGAGTTTCCAGTTAGAG 595  
 QY 186 LeuGlyValIleSerIleThrPheGlnLeuSerSerHisProIleLeuGlyAspTyrPser 205  
 |||||  
 |||||  
 Db GGTGGCTCAAGCAATTTTCTTTTCCCTCATCATCAGAGCCCTTCCAGGGCTTCTCAAG 655  
 QY 206 IleGlnValGlnValAsnAsp-----GlnThrTyrTyrGlnSerPheGlnValSerGlu 223  
 |||||  
 |||||  
 Db GTGGGTGTCACAGAAATCAGGTGAGAGAGACAGAGAC--CTTTTCACTGAGAGAA 712  
 QY 224 TyrValLeuProLysPheGlnValThrLeuGlnThrProLeuTyrCysSerMetAsnSer 243  
 |||||  
 |||||  
 Db TTGTGTCCTCCCAAGTTGATGATGACAGTAACAGTCCAAAGATATCATCATCTTGAA 772  
 QY 244 LysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProValIleGlyAsp 263  
 |||||  
 |||||  
 Db GAAGAGATGAATGTATCAGTGTGTGCTTATACATATGAGAGAGCTGTCTCTGAGACAT 832  
 QY 264 ValThrLeuThrPheLeuProLeu----- 271  
 |||||  
 |||||  
 Db GTGACTGTGAGCATTTGGCAAGAAATGATGACGCTTCCGACTGCCAGGTGAAGATTCA 892  
 QY 272 ---SerPheTyrGlyLys 276  
 |||||  
 |||||  
 Db CAGGCTTTCTGTGAGAAATTCAGTGAACGCTAAACAGCATGCGCTTATATACGAA 952  
 QY 277 LysIleAsnIleThrIleThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAsp 296  
 |||||  
 |||||  
 Db GTAATA-----ACCAAGGTCTTCCAGTGAAGAG-----AAGGAG 988

QY 297 GluGluMetLysAsnValMetAspSerSerAsnGlyLeuSerGluTyrIleAspLeuSer 316  
 |||||  
 |||||  
 Db TATGAATGAA-----CTTCACTAGAGGCCAGATTCOA 1024  
 QY 317 SerProGlyProValGluIleLeuThr-----ThrValThrGluSerVal 331  
 |||||  
 |||||  
 Db GAAGAGAGAAAGTGTGGAATGTAGTGAAGGAGTCCAGTCAATGAATCAAGAACATTA 1084  
 QY 332 ThrGlyIleSerArgAsnValSerThrAsnValPhePheGlyGlnHisAspTyrIleIle 351  
 |||||  
 |||||  
 Db ACCAACTCTCA---TTTGTGAAGTGAAGTCACTTTCAGACAGGA-----ATT 1132  
 QY 352 GluPhePheAspTyrThrThrValLeuLysProSerLeuAsnPheThrAlaThrValLys 371  
 |||||  
 |||||  
 Db CCTTCTTTGGG-----CAG 1147  
 QY 372 ValThrArgAlaAspGlyAsnGlnLeuThrLeuGluAlaArgArgAsnValIle 391  
 |||||  
 |||||  
 Db GTGGCTTGTAGTGAATGGAAGGCGTCCCTATACCA-----AATGAATCATTA 1195  
 QY 392 ThrValThrGlnArgAsnTyrThrGluTyrTyrPserGlySerAsnSerGlyAsnGlnLys 411  
 |||||  
 |||||  
 Db TTCACT--AGAGAAATGAGCAAACTATTAC-----TCCAAATGCTACCGAGATGAG 1246  
 QY 412 MetGluAlaValGln---LysIleAsnTyrThr-----ValProGlnSerGlyThrPhe 428  
 |||||  
 |||||  
 Db CATGGCTTGTATCATTTCTATATCAACCAACCAAGTTATGGATACCTCTTACTGT 1306  
 QY 429 LysIleGluPhePro-----IleLeuGluAspSer 438  
 |||||  
 |||||  
 Db AGGCTCAATTACAAAGATGATGCTTCCCTTACGCGTCAAGTGGGTGTCAGAAACACAC 1366  
 QY 439 SerGluLeuGlnLeuLysAlaIlePheLeuGlySerLysSerSerMetAlaValHisSer 458  
 |||||  
 |||||  
 Db GAAGAGCATCATCATCTTAT----- 1390  
 QY 459 LeuPheLysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLys 478  
 |||||  
 |||||  
 Db CTGTGTTCTCCCAAGCAAGAGCTTGTCCACTTGAAGCCCATGCTCATGAATCACTACC 1450  
 QY 479 ValGlySer-----ProPheGluLeuValIleSerGlyAsnLys-----Arg 492  
 |||||  
 |||||  
 Db TGTGCTCATCTCAACAGTCCAGGACATATATTGTAAGAGAGGACCTGCTGGGG 1510  
 QY 493 LeuLysGluLeuSer-----TyrMetValValSerArgGlyGlnLeuValAlaValGly 510  
 |||||  
 |||||  
 Db CTGAAGAGCTCTCTTATTATCTGATTAATGGCAAGGAGGAGCATTTGCGAACTGGG 1570  
 QY 511 -----LysGlnAsnSerThr-----MetPheSerLeuThr 520  
 |||||  
 |||||  
 Db ACTCATGACTCTGTGTAAGCAGAGAAACATGAAGGCGCATTTTCCATCTCAATTCCT 1630  
 QY 521 ProGluAsnSerTyrThrProLysAlaCysValIleValTyrTyrIleGluAspAspGly 540  
 |||||  
 |||||  
 Db GTGAAGTCAAGATGCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1690  
 QY 541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560  
 |||||  
 |||||  
 Db GACGTGATGGGATTCGCAAAATATGATGTGAAATTTGTCTGCGCAACAGAGGTGAT 1750  
 QY 561 LeuTyrTyrPserLysValIleValGluProSerGluLysValSerLeuArgIleSerVal 580  
 |||||  
 |||||  
 Db TTGACTTCAAGCCATCAACAAATGCTCCAGCTTACACGCCACCTGCGATCAAGG 1810  
 QY 581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600  
 |||||  
 |||||  
 Db GCT---CCTCAGTCTGCTGCGCCCTCGGTGTGTGAGCAAAAGCTGCTGCTATGAAG 1867  
 QY 601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeu---GluLeuTyrAsnThr 619  
 |||||  
 |||||  
 Db CCTGATGCTGAGCTCTCGCGCTCGGTTTACAACTGCTTACAGAAAGAGACCTTCACT 1927  
 QY 620 GlyTyr-----TyrLeuGlyMet 625



1928 GGGCTCCCTGGGCTTTGAATGACGAGCAGTGAAGTGCATCAATCGTCAATATGTC 1987  
Qy 626 PheMetAnser-----PheAla 631  
Db 1988 TATTTAATGGAATCACAATATCTCCAGTATCAAGTCAAAATGAAAGATATGTACAGC 2047  
Qy 632 ValPheGlnGlySerGlyLeuTrpValLeuThrAspAlaLeuLeuThrLysAspTrpLe 651  
Db 2048 TTCTAGAGGACATGGCTTAAAGGACATTCACCAACTCAAAAGATTCGTAAACCAAAATG 2107  
Qy 652 AspGlyValTyraAspAlaGln-----Tyr 660  
Db 2108 TGTCCACAGCTTCAACAGTAAATGATGACCTGAGCTCACTGAGCTTTTAT 2167  
Qy 661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680  
Db 2168 GAGTCAGATGTAATGGGAAGAGCCCATGACCGCTGGTGCATGTGAAG----- 2218  
Qy 681 GlySerSerProHis-----ValArgLysHisPheProGluTrpIleTrpLeu 697  
Db 2219 -----CCTCACAGGAGAACCGTACGAAAGTACTTCCCTGAGACATGATCGGAT 2269  
Qy 698 AspThrAsnMetGlyTyraGlyIleTyrglnGluPheGluValThrValProAspSerIle 717  
Db 2270 TTGGTGGTGAATCACTACAGCAGGGGTGGCTGAGTAAAGTAAACAGTCCCTGACACCATC 2329  
Qy 718 ThrSerTrpValAlaThrGlyPheValIleSerGluAsnGlyLeuGlyLeuThr 737  
Db 2330 ACCAGTGAAGAGGAGGGGCTTCTGCTGTGAAGATGTGACTGTGATCTCTTC 2389  
Qy 738 ThrProValGluLeuGlnAlaPheGlnProPheIlePheLeuAsnLeuProTySer 757  
Db 2390 ACT---GCTCTCTCCAGCCTTCCAGCCTTCTTGTGAGCTTACAAAGCCTTACTCT 2446  
Qy 758 ValIleArgGlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrlLeuLysAspAla 777  
Db 2447 GTGATTCGTGGAGAGGCTTCACTACACTCAAGCCAGCTCCAACTACCTCCCAAAATGC 2506  
Qy 778 ThrIleValLysValIleGlnLysSerAspLysPheAspIleLeuMetThrSer 797  
Db 2507 ATCCGGTTCAGTGTGAGCTGAAGCCCTCCGCTCTCTGCTGCCAGTGGAGAG 2566  
Qy 798 Gln-----IleAsnAlaThrGlyHisGlnGlnThrLeuValProSer 812  
Db 2567 GAACAAGCGCTCACTGCTATGTGCAACGGGGCGCA----- 2605  
Qy 813 GluAspGlyAlaThrValLeuPheProIleArgProThrHisIleGlyGluIleProIle 832  
Db 2606 -----ACTGTGCTCTGGGAGTAAACCCAAAGTCAATTAGAAATGTGAATTC 2653  
Qy 833 ThrValThrAlaLeuSer-----ProThrAla 841  
Db 2654 ACTGTGAGCGAGGACACTAGACTCAAGAGCTGTGTGGAGCTGAGGCTTCAAGTT 2713  
Qy 842 -----SerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIle 856  
Db 2714 CTTGAACACGGAAGAAAGACACAGTCAAGCTCTGTTGGTGAACCTGAAAGGACTA 2773  
Qy 857 GluLysSerTySerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThr 876  
Db 2774 GAGAAAGAAACAACTTCACTCCCTTGTCTCATCAGTGTGAGTTCT----- 2827  
Qy 877 LeuLysThrLeuSerPheSerPheProProAsnThrValThrGlySerGluArgValGln 896  
Db 2828 ---GAACAATATTCCTCGAAACTGCAACCAATGTGTAGAAAGATCTGCCGAGCTTCT 2884  
Qy 897 IleThrAlaIleGlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArg 916  
Db 2885 GTCTCAGTTTGGAGACATATTAAGCTCTGCCATGCAAAACACACAAATCTTCTCAG 2944  
Qy 917 MetProTyrglyCysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrlLeu 936  
|||:::  
1928 GGGCTCCCTGGGCTTTGAATGACGAGCAGTGAAGTGCATCAATCGTCAATATGTC 1987  
Qy 626 PheMetAnser-----PheAla 631  
Db 1988 TATTTAATGGAATCACAATATCTCCAGTATCAAGTCAAAATGAAAGATATGTACAGC 2047  
Qy 632 ValPheGlnGlySerGlyLeuTrpValLeuThrAspAlaLeuLeuThrLysAspTrpLe 651  
Db 2048 TTCTAGAGGACATGGCTTAAAGGACATTCACCAACTCAAAAGATTCGTAAACCAAAATG 2107  
Qy 652 AspGlyValTyraAspAlaGln-----Tyr 660  
Db 2108 TGTCCACAGCTTCAACAGTAAATGATGACCTGAGCTCACTGAGCTTTTAT 2167  
Qy 661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680  
Db 2168 GAGTCAGATGTAATGGGAAGAGCCCATGACCGCTGGTGCATGTGAAG----- 2218  
Qy 681 GlySerSerProHis-----ValArgLysHisPheProGluTrpIleTrpLeu 697  
Db 2219 -----CCTCACAGGAGAACCGTACGAAAGTACTTCCCTGAGACATGATCGGAT 2269  
Qy 698 AspThrAsnMetGlyTyraGlyIleTyrglnGluPheGluValThrValProAspSerIle 717  
Db 2270 TTGGTGGTGAATCACTACAGCAGGGGTGGCTGAGTAAAGTAAACAGTCCCTGACACCATC 2329  
Qy 718 ThrSerTrpValAlaThrGlyPheValIleSerGluAsnGlyLeuGlyLeuThr 737  
Db 2330 ACCAGTGAAGAGGAGGGGCTTCTGCTGTGAAGATGTGACTGTGATCTCTTC 2389  
Qy 738 ThrProValGluLeuGlnAlaPheGlnProPheIlePheLeuAsnLeuProTySer 757  
Db 2390 ACT---GCTCTCTCCAGCCTTCCAGCCTTCTTGTGAGCTTACAAAGCCTTACTCT 2446  
Qy 758 ValIleArgGlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrlLeuLysAspAla 777  
Db 2447 GTGATTCGTGGAGAGGCTTCACTACACTCAAGCCAGCTCCAACTACCTCCCAAAATGC 2506  
Qy 778 ThrIleValLysValIleGlnLysSerAspLysPheAspIleLeuMetThrSer 797  
Db 2507 ATCCGGTTCAGTGTGAGCTGAAGCCCTCCGCTCTCTGCTGCCAGTGGAGAG 2566  
Qy 798 Gln-----IleAsnAlaThrGlyHisGlnGlnThrLeuValProSer 812  
Db 2567 GAACAAGCGCTCACTGCTATGTGCAACGGGGCGCA----- 2605  
Qy 813 GluAspGlyAlaThrValLeuPheProIleArgProThrHisIleGlyGluIleProIle 832  
Db 2606 -----ACTGTGCTCTGGGAGTAAACCCAAAGTCAATTAGAAATGTGAATTC 2653  
Qy 833 ThrValThrAlaLeuSer-----ProThrAla 841  
Db 2654 ACTGTGAGCGAGGACACTAGACTCAAGAGCTGTGTGGAGCTGAGGCTTCAAGTT 2713  
Qy 842 -----SerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIle 856  
Db 2714 CTTGAACACGGAAGAAAGACACAGTCAAGCTCTGTTGGTGAACCTGAAAGGACTA 2773  
Qy 857 GluLysSerTySerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThr 876  
Db 2774 GAGAAAGAAACAACTTCACTCCCTTGTCTCATCAGTGTGAGTTCT----- 2827  
Qy 877 LeuLysThrLeuSerPheSerPheProProAsnThrValThrGlySerGluArgValGln 896  
Db 2828 ---GAACAATATTCCTCGAAACTGCAACCAATGTGTAGAAAGATCTGCCGAGCTTCT 2884  
Qy 897 IleThrAlaIleGlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArg 916  
Db 2885 GTCTCAGTTTGGAGACATATTAAGCTCTGCCATGCAAAACACACAAATCTTCTCAG 2944  
Qy 917 MetProTyrglyCysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrlLeu 936  
|||:::  
2945 ATGCCCTATGCTGTGGAGAGCAGAAATATGCTCTTGTGCTTCAACATCTATGACTG 3004  
Qy 937 AspTyrlLeuThrLysLysGlnLeuThrAspAsnLeuLysGluValAlaLeuSerPhe 956  
Db 3005 GATTATCTAATATGAACACAGCAGCTTATCTCCAGAGTCAAGTCCAAAGGCAATGGCTAT 3064  
Qy 957 MetArgGlnGlyTyrglnArgGluLeuLeuTyrglnArgGluAspGlySerPheSerAla 976  
Db 3065 CTCACACTGTTTACCAGAGACAGTGAACCTAACAACATATATGATGGCTCTACAGCACC 3124  
Qy 977 PheGly---AsnTyraAspProSer---GlySerThrTrpLeuSerAlaPheValLeuArg 994  
Db 3125 TTGGGAGAGGAAATGACAGAAACAGGAGCAACCTGGCTGCACACCTTGTGTGAAAG 3184  
Qy 995 CysPheLeuGluAlaAspProTyrlIleAspIleAspGlnAsnValLeuHisArgTrpTy 1014  
Db 3185 ACTTTGGCCCAAGCTGAGGCTTACATCTTCATGATGAAACACACATTCACCAAGCCCTC 3244  
Qy 1015 ThrTrpLeuLysGlyHisGlnLysSerAsnGlyGluPheTrpAspProGlyArgValIle 1034  
Db 3245 ATATGGCTCCAGAGGAGGAGAAAGCAATGGCTTTTCAGAGCTCTGGGTCACTGCTC 3304  
Qy 1035 HisSerGluLeuGlnGlyGlyAsnLysSerProValThrLeuThrAlaTyrlIleValThr 1054  
Db 3305 AACAAATGCCATAAAGGAGAGTGAAGATGAAGTACCTTCGCGCTATATCACCATC 3364  
Qy 1055 SerLeuLeuGlyTyraArgLysTyrglnProAsnIleAspValGlnGluSerIleHisPhe 1074  
Db 3365 GCCCTTGTGAGATTCCTCTCACAGTCACTACCTCTGTGTGCGCAATGCCCTGTTTGC 3424  
Qy 1075 LeuGluSer-----GluPheSerArgGlyLysSerAspAsnTyrlThr 1088  
Db 3425 CTGAGTCAAGCCTCGAAGACAGCAAGAAAGGAGCAATAGC---AGCCATGATATATAC 3481  
Qy 1089 LeuAlaLeuIleThrTyrlAlaLeuSerSerValGlySer---ProLysAlaLysGluAla 1107  
Db 3482 AAAGCACTGTGCTCATGTGCTTGGCTGCGCAGGTAAACAGGACAAAGAGAAAGTA 3541  
Qy 1108 LeuAsnMetLeuThrTrpArgIleGlnGlnGlyGlyMetGlnPheTrp----- 1124  
Db 3542 CTCAAGTCACTTATAGAGAGCTGTGAAGAAAGACAACTCTCTCCATTGGAGCGCCCT 3601  
Qy 1125 ValSerSerGlnLysLeuSerAspSerTrpGlnProArg-----SerLeuAspIle 1142  
Db 3602 CAGAAACCAAGCAGCAGAGGGGCAATTTTACAGAACCCAGGCTCCCTGCTGAGGTG 3661  
Qy 1143 GluValAlaAlaTyrlAlaLeuLeuSerHisPheLeuGlnPheGln-----ThrSer 1159  
Db 3662 GAGATGACATCTTATGTGCTCTCGCT---TATCTCACGCGCCAGCCAGCCCAACCTCG 3718  
Qy 1160 GluGlyIlePro-----IleMetArgTrpLeuSerArgGlnArgAsnSerLeu 1175  
Db 3719 GAGGAGCTGAGCTCTGCAACCAACATCGTGAAGTGAATCAAGAGCAGCAAGATGCGCAG 3778  
Qy 1176 GlyGlyPheAlaSerThrGlnAspThrThrValAlaLeuLysValAlaLeuSerGluPheAla 1195  
Db 3779 GGGGTTTCTCCCTCCACAGCAGACAGATGTGGCTCTCAATGCTCTGTGCAAAATATGGA 3838  
Qy 1196 AlaLeuMetAsnThrGluArgThrAsnIleGlnValThrValThrGlyProSerSerPro 1215  
Db 3839 GCCGCCACATTTTACC---AGGACTGGGAAGGCTGACAGGTGACTATCCAGTCTTCAGGG 3895  
Qy 1216 SerProLeuAlaValAlaGln----- 1222  
Db 3896 ACAATTTCCAGCAAAATTCCAAGTGAACAACAATGCGCTTACTGACAGAGGCTCTCA 3955  
Qy 1223 -----ProMetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCys 1238  
Db 3956 TTGCCAGAGCTGCTGGGAATACAGCATGAAGTAAAGTACAGAGAAAGATGTCTACCTC 4015  
Qy 1239 GlnLeuAsnValValTyraAsnValLysAlaSerGlySerSerArgArgArgSerIle 1258  
|||:::  
4016 CAGACCTCTTGAAATATACATATT-----CTC 4042

2945 ATGCCCTATGCTGTGGAGAGCAGAAATATGCTCTTGTGCTTCAACATCTATGACTG 3004  
Qy 937 AspTyrlLeuThrLysLysGlnLeuThrAspAsnLeuLysGluValAlaLeuSerPhe 956  
Db 3005 GATTATCTAATATGAACACAGCAGCTTATCTCCAGAGTCAAGTCCAAAGGCAATGGCTAT 3064  
Qy 957 MetArgGlnGlyTyrglnArgGluLeuLeuTyrglnArgGluAspGlySerPheSerAla 976  
Db 3065 CTCACACTGTTTACCAGAGACAGTGAACCTAACAACATATATGATGGCTCTACAGCACC 3124  
Qy 977 PheGly---AsnTyraAspProSer---GlySerThrTrpLeuSerAlaPheValLeuArg 994  
Db 3125 TTGGGAGAGGAAATGACAGAAACAGGAGCAACCTGGCTGCACACCTTGTGTGAAAG 3184  
Qy 995 CysPheLeuGluAlaAspProTyrlIleAspIleAspGlnAsnValLeuHisArgTrpTy 1014  
Db 3185 ACTTTGGCCCAAGCTGAGGCTTACATCTTCATGATGAAACACACATTCACCAAGCCCTC 3244  
Qy 1015 ThrTrpLeuLysGlyHisGlnLysSerAsnGlyGluPheTrpAspProGlyArgValIle 1034  
Db 3245 ATATGGCTCCAGAGGAGGAGAAAGCAATGGCTTTTCAGAGCTCTGGGTCACTGCTC 3304  
Qy 1035 HisSerGluLeuGlnGlyGlyAsnLysSerProValThrLeuThrAlaTyrlIleValThr 1054  
Db 3305 AACAAATGCCATAAAGGAGAGTGAAGATGAAGTACCTTCGCGCTATATCACCATC 3364  
Qy 1055 SerLeuLeuGlyTyraArgLysTyrglnProAsnIleAspValGlnGluSerIleHisPhe 1074  
Db 3365 GCCCTTGTGAGATTCCTCTCACAGTCACTACCTCTGTGTGCGCAATGCCCTGTTTGC 3424  
Qy 1075 LeuGluSer-----GluPheSerArgGlyLysSerAspAsnTyrlThr 1088  
Db 3425 CTGAGTCAAGCCTCGAAGACAGCAAGAAAGGAGCAATAGC---AGCCATGATATATAC 3481  
Qy 1089 LeuAlaLeuIleThrTyrlAlaLeuSerSerValGlySer---ProLysAlaLysGluAla 1107  
Db 3482 AAAGCACTGTGCTCATGTGCTTGGCTGCGCAGGTAAACAGGACAAAGAGAAAGTA 3541  
Qy 1108 LeuAsnMetLeuThrTrpArgIleGlnGlnGlyGlyMetGlnPheTrp----- 1124  
Db 3542 CTCAAGTCACTTATAGAGAGCTGTGAAGAAAGACAACTCTCTCCATTGGAGCGCCCT 3601  
Qy 1125 ValSerSerGlnLysLeuSerAspSerTrpGlnProArg-----SerLeuAspIle 1142  
Db 3602 CAGAAACCAAGCAGCAGAGGGGCAATTTTACAGAACCCAGGCTCCCTGCTGAGGTG 3661  
Qy 1143 GluValAlaAlaTyrlAlaLeuLeuSerHisPheLeuGlnPheGln-----ThrSer 1159  
Db 3662 GAGATGACATCTTATGTGCTCTCGCT---TATCTCACGCGCCAGCCAGCCCAACCTCG 3718  
Qy 1160 GluGlyIlePro-----IleMetArgTrpLeuSerArgGlnArgAsnSerLeu 1175  
Db 3719 GAGGAGCTGAGCTCTGCAACCAACATCGTGAAGTGAATCAAGAGCAGCAAGATGCGCAG 3778  
Qy 1176 GlyGlyPheAlaSerThrGlnAspThrThrValAlaLeuLysValAlaLeuSerGluPheAla 1195  
Db 3779 GGGGTTTCTCCCTCCACAGCAGACAGATGTGGCTCTCAATGCTCTGTGCAAAATATGGA 3838  
Qy 1196 AlaLeuMetAsnThrGluArgThrAsnIleGlnValThrValThrGlyProSerSerPro 1215  
Db 3839 GCCGCCACATTTTACC---AGGACTGGGAAGGCTGACAGGTGACTATCCAGTCTTCAGGG 3895  
Qy 1216 SerProLeuAlaValAlaGln----- 1222  
Db 3896 ACAATTTCCAGCAAAATTCCAAGTGAACAACAATGCGCTTACTGACAGAGGCTCTCA 3955  
Qy 1223 -----ProMetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCys 1238  
Db 3956 TTGCCAGAGCTGCTGGGAATACAGCATGAAGTAAAGTACAGAGAAAGATGTCTACCTC 4015  
Qy 1239 GlnLeuAsnValValTyraAsnValLysAlaSerGlySerSerArgArgArgSerIle 1258  
|||:::  
4016 CAGACCTCTTGAAATATACATATT-----CTC 4042



Db 1148 GTGGCCCTAGATGGAGGCGTCCATACCA-----AATAAGCATTA 1195  
 Qy 392 ThrValThrGlnArgAsnThrGluTyrTrpSerGlySerAsnSerGlyAsnGlnLys 411  
 Db 1196 TTTCATC---AGAGGAATGAAGCAAACTATTAC-----TCCATGCTCACCACGAGAG 1246  
 Qy 412 MetGlnAlaValGln---LysIleAsnTyrThr-----ValProGlnSerGlyThrPhe 428  
 Db 1247 CATGCGCTTGATCAGATTCTCTATCAACACCAACGTTATGGTACTCTTACTGTT 1306  
 Qy 429 LysIleGluPhePro-----LleuGlnAspSer 438  
 Db 1307 AGGGTCATTACCAAGATCGATGCCCTTGTAAGGCTTACCAAGTGCTCGAAGAACAC 1366  
 Qy 439 SerGluLeuGlnLeuLysAlaTyrPheLeuGlySerSerSerMetAlaValHisSer 458  
 Db 1367 GAAGAGGACATCACACCTGTTAT----- 1390  
 Qy 459 LeuPheLysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGlnAsnIleLys 478  
 Db 1391 CTGTGTTCTCCCAAGACAGAGCTTGTCCACCTTGACCCATGTCATGAATACCC 1450  
 Qy 479 ValGlySer-----ProPheGluLeuValValSerGlyAsnLys-----Arg 492  
 Db 1451 TGTGGCCATCTCAGACAGTCACAGGCACATTATATCTGAATGAGGACCCCTGCGGG 1510  
 Qy 493 LeuLysGluLeuSer-----TyrMetValValSerArgGlyGlnLeuValAlaValGly 510  
 Db 1511 CTGAAGAACTCTCTTTTATTATCTGATTAATGGCAAGGAGGACATTGTCGAACCTGG 1570  
 Qy 511 -----LysGlnAsnSerThr-----MetPheSerLeuThr 520  
 Db 1571 ACTCATGAGCTGCTGTGACAGAGAGACATGAAGGCGCATTTTTCATCTCATCTCCCT 1630  
 Qy 521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGlnAspAspGly 540  
 Db 1631 GTGAGGACGACATTTGCTCTCGTCCGTGCTGCTCATCTATGCTGTTTACCTAACCGGG 1690  
 Qy 541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560  
 Db 1691 GACGTGATGGGAGATTCTGCAGAAATATGATGTTGAAATTTGTGCGCCACAGAGTGAT 1750  
 Qy 561 LeuTyrTrpSerLysValLysAlaGluProSerGlyLysValSerLeuArgIleSerVal 580  
 Db 1751 TTGAGCTTCACGCCATCACAAGTCTCCAGCTCCACAGCCACCTCGAGTGCACAGCG 1810  
 Qy 581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600  
 Db 1811 GCT---CCTCAGTCCGTCTGCGCCCTCGTCTGTGACCAAGGCTGCTCATGAG 1867  
 Qy 601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeu---GluLeuTyrAsnThr 619  
 Db 1868 CCGATGCTGAGCTCTCGCGCTCTCGGTTTACACCTGCTACCAAGAAAGACCTCACT 1927  
 Qy 620 GlyTyr-----TyrLeuGlyMet 625  
 Db 1928 GCGTTCCTCGGCGCTTGAATGACCAAGACGATGAAGCTGCATCAATCGTCAATATGTC 1987  
 Qy 626 PheMetAsnSer-----PheAla 631  
 Db 1988 TATATTATGCAATCACAATATCTCACTCACTCACTCACTCACTCACTCACTCACTCACT 2047  
 Qy 632 ValPheGlnGluCysGlyLeuTyrValLeuThrAspAlaAsnLeuThrLysAspTyrIle 651  
 Db 2048 TTCTTAGAGGACATGGGCTTAAAGGCACTTACCAACTCAAGATTCTGTTAAACCCCAATG 2107  
 Qy 652 AspGlyValTyrAspAsnAlaGlu-----Tyr 660  
 Db 2108 TGTCCACAGCTTCACACATATGAATGCATGAGCTGAAAGCTTACGCTGAGTTTAT 2167  
 Qy 661 AlaGluArgPheMetGluGlnAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680  
 Db 2168 GAGTCAGATGTAATGGAAGAGGCAATGACGCTGTGATGTTGAAGAG----- 2218

Qy 681 GlySerSerProHis-----ValArgLysHisPheProGluThrTrpIleTrpLeu 697  
 Db 2219 -----CCTCACACGAGACCGTATGCAAAAGTACTCTCCATGACATGATCTGGAT 2269  
 Qy 698 AspThrAsnMetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIle 717  
 Db 2270 TTGGTGTGTAACTCAGACAGGGGTGGCTGAGTGAAGTAAACAGTCCCTGACACATC 2329  
 Qy 718 ThrSerTrpValAlaThrGlyPheValIleSerGluAsnLeuGlyLeuGlyLeuThrThr 737  
 Db 2330 ACCGAGTGAAGGACAGGGGCTTCTGCTGCTCTGAGATCTGAGACTTGTATCTCTCC 2389  
 Qy 738 ThrProValGluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSer 757  
 Db 2390 ACT---GCTCTCTCCGAGGCTTCCAGCCCTCTTGTGAGACTTACAAATGCTTACTCT 2446  
 Qy 758 ValIleArgGlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAla 777  
 Db 2447 GTGATTCGTGAGAGGCTTCCACACTCAGCCACGCTCTTAAACTACTCCCAATGC 2506  
 Qy 778 ThrGluValLysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSer 797  
 Db 2507 ATCCGGGTCAGTGACGCTGAGAGCTCTCCGCTTCTTGCTGCTCCAGTGGAGAG 2566  
 Qy 798 Glu-----IleAsnAlaThrGlyHisGlnGlnThrLeuLeuValProSer 812  
 Db 2567 GAACAAAGGCTCAGTCACTGCTGCAAGGCGGCA----- 2605  
 Qy 813 GluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGlyGluIleProIle 832  
 Db 2606 -----ACTGTCTCTGGCGGAGTAAACCCAAAGTCATTAGCAAAATGCAATTC 2653  
 Qy 833 ThrValThrAlaLeuSer-----ProThrAla 841  
 Db 2654 ACTGTAGGCGACAGGACACTAGAGTCAAAGCTGTGTGGACTGAGTGCCTTCACTT 2713  
 Qy 842 -----SerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIle 856  
 Db 2714 CTGAACACGGAAGAAAGACAGTCAATCAGAGCTCTGTTGGTGAACCTGAAGAGCTA 2773  
 Qy 857 GluLysSerTyrSerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThr 876  
 Db 2774 GAGAAAGAAACCAATCACTCCCTTCTTGTCCATCAGAGTGTGAGTTCCTTCTT 2827  
 Qy 877 LeuLysThrLeuSerPheSerPheProAsnThrValThrGlySerGluArgValGln 896  
 Db 2828 ---GAAGAATTAATCTGAAATCTGCCCAAAATGTGTAGAAGATCTGCCGAGCTTCT 2884  
 Qy 897 IleThrAlaIleGlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArg 916  
 Db 2885 GTTCACATTTTGGAGAGCATTTATAGCTCTGCCATGCAAAACACAAATCTTCTCAG 2944  
 Qy 917 MetProTyrGlyCysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeu 936  
 Db 2945 ATGCCATATGCTGTGAGAGAGCAAAATATAGTCTCTTCTCTCAATCACTATGTACTG 3004  
 Qy 937 AspTyrLeuThrLysLysGluLeuThrAspAsnLeuLysGluLysAlaLeuSerPhe 956  
 Db 3005 GATTATCTTAATGAAMAACAGACAGCTTACTCAGAGGTCAAGTCAAGGCGCATTGCTAT 3064  
 Qy 957 MetArgGlnGlyTyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAla 976  
 Db 3065 CTCAACACTGTTTCCAGACAGCTTGAATCACTCAAAACATATATGCTCTCTACAGACAC 3124  
 Qy 977 PheGly---AsnTyrAspProSer---GlySerThrTrpLeuSerAlaPheValLeuArg 994  
 Db 3125 TTTCGGAGAGGATATGACAGAAACAGACACCTGGCTTCTTGTCTGAGAG 3184  
 Qy 995 CysPheLeuGluAlaAspProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyr 1014  
 Db 3185 ACTTTGCCCAAGCTCGAGGCTTCACTCTCATGATGAAAGACACATTAACCAAGCCCTC 3244

QY 1015 ThrTrpLeuLysGlyHISGlnLysSerAsnGlyValPheTrpAspProGlyValVal1110 1034  
 DB 3245 ATATGCTCTCCAGAGAGAGAGCAATGCTGTTCCAGAGCTCTGCTGCTACCTCC 3304  
 QY 1035 HisSerGluLeuGlnGlyValAsnLysSerProValThrLeuPheValTyrIleValThr 1054  
 DB 3305 AACATGCTTAAAGAGAGAGAGTGAAGATGAAGTAACTCTCCGCTTATATACCAATC 3364  
 QY 1055 SerLeuLeuGlyTyrTrpGlyTrpGlnProAsnIleAspValGlnLysSerIleHisPhe 1074  
 DB 3365 GCCCTTCTGGAAGATTCCTCTCAAGTCACCTCACCTGTTGCGCAATGCGCTTTTTC 3424  
 QY 1075 LeuGluSer-----GluPheSerArgIlyIleSerAspAntyTrp 1088  
 DB 3425 CTGGAGCTGAGCTCGAAGACAGACCAAGAGGAGCAATGGC---AGCCATGATATATACC 3481  
 QY 1089 LeuValLeuIleThrTrpAlaLeuSerSerValGlySer---ProIysAlaIysGluAla 1107  
 DB 3482 AAGACAGCTGCTGCTTATGCTTTTGGCTTGGCAGGTAAACAGAGACAGAGAGAGTA 3541  
 QY 1108 LeuAsnMetLeuThrTrpArgAlaGlnGlnGlyIleGlyMetGlnPheTrp----- 1124  
 DB 3542 CTCAATCACTTATATAGAGAGAGCTGGAAGAAAGACACTCTGCTCATTCGGAGCGCCCT 3601  
 QY 1125 ValSerSerGluSerLysLeuSerAspSerTrpGlnProArg-----SerLeuAspIle 1142  
 DB 3602 CAGAAACCAAGGACAGCAGGAGGAGCATTTTTCAGAACCCAGAGCTCCCTGCTGAGG 3661  
 QY 1143 GluValAlaAlaTrpAlaLeuLeuSerHisPheLeuGlnPheGln-----ThrSer 1159  
 DB 3662 GAGATGACATCTTATGCTGCTCTGCTGCT---TATCTACGAGCCAGGAGCCCAACCTCG 3718  
 QY 1160 GluGlyIlePro-----IleMetArgTrpLeuSerArgIleArgAsnSerLeu 1175  
 DB 3719 GAGGAGCTGACCTCTGCAACCAACATCGTGAAGTGAATCAGAACAGACAGATGCTCCAG 3778  
 QY 1176 GlyGlyPheAlaSerThrGlnAspThrTrpValAlaLeuValAsnLysSerGluPheAla 1195  
 DB 3779 GGGGCTTCTCTCCACCCAGAGACAGAGTGGTCTCTCATGCTCTGCTCAAAATATGAG 3838  
 QY 1196 AlaLeuMetAsnThrGluArgThrAsnIleGlnValThrValThrGlyProSerSerPro 1215  
 DB 3839 GCGGCACATTTTACC---AGGACTGGAGAGGCTGCAAGGTGACTATTCACAGTCTTCAGG 3895  
 QY 1216 SerProLeuAlaValValGln----- 1222  
 DB 3896 ACATTTTCCAGCAATTCAGAGTGGACAAACAAATCGCTTACTGACAGAGTCTCA 3955  
 QY 1223 -----PrometAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCys 1238  
 DB 3956 TTGCCAGAGCTGCTGGGGAATACAGCATGAAGTGAAGAGAGATGTGTCTACCTC 4015  
 QY 1239 GlnLeuAsnValValIlyTrpAsnValIlybAlaSerGlySerSerArgArgArgSerIle 1258  
 DB 4016 CAGACCTCTTGAATAACATATAT-----CTC 4042  
 QY 1259 GlnAsnGlnGluAlaPheAspLeuAspValAlaValLys-----GluAsnLysAsp 1275  
 DB 4043 CCAAGAAAGAGAGATTCCTGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 4102  
 QY 1276 AspLeuAsn---HisValAspLeuAsnValCysThrSerPheSerGlyProGly----- 1292  
 DB 4103 GAACCAAAAGCCACACACAGCTTCCAAATCTCCCTAAGTGTCTACACAGAGAGACCC 4162  
 QY 1293 ---ArgSerGlyMetAlaLeuMetGluValAsnLeuLysSerGlyPheMetValProSer 1311  
 DB 4163 TCTGCTCCAAACATGCGATCTGTGAAGTGAAGATGTCTCTGCTTCAATTCCTCCTGA 4222  
 QY 1312 GluAlaIleSerLeu-----SerGluThrValLysValGluValTrpAspHisGly 1328  
 DB 4223 CCAACAGTGAAGATCTTGAAGATCTAATCATATGAGCCGACAGACAGTCAAGCAAC 4282  
 QY 1329 LysLeuAsnLeuTrpLeuAspSerValAsnGluThrGlnPheCysValAsnIleProAla 1348

DB 4283 CATGCTTGATTACTTATGATTAAGGTTCAAATTCAGACACTGAGCTTGTCTTACCGTT 4342  
 QY 1349 ValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSerIleValAspTrp 1368  
 DB 4343 CTGCAAGATGTCCTCCAGTAAAGATCTCAACCAAGCATATGATTAATTAATTAAT 4402  
 QY 1369 GluProArgArgGlnAlaValArgSerTrpAsnSer 1380  
 DB 4403 GAGAGGAGATGATTTGCAATGCTGATGATTAATGCT 4438  
 RESULT 11  
 US-10-292-081A-3  
 ; Sequence 3, Application US/10292081A  
 ; Publication No. US20030162202A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kenneth David Becker  
 ; APPLICANT: Gonul Velicelebi  
 ; APPLICANT: Xin Wang  
 ; APPLICANT: Randolph E. Tanzi  
 ; APPLICANT: Lars Berttram  
 ; APPLICANT: Aleister J. Saunders  
 ; TITLE OF INVENTION: SINGLE NUCLEOTIDE POLYMORPHISMS AND MUTATIONS ON ALPHA-2-MACROG  
 ; FILE REFERENCE: 37481-3323  
 ; CURRENT APPLICATION NUMBER: US/10/292,081A  
 ; CURRENT FILING DATE: 2002-11-08  
 ; PRIOR APPLICATION NUMBER: 60/337434  
 ; PRIOR FILING DATE: 2001-11-09  
 ; NUMBER OF SEQ ID NOS: 15  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 3  
 ; LENGTH: 4577  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-292-081A-3  
 Alignment Scores:  
 Pred. No.: 3,566-133 Length: 4577  
 Score: 1440.00 Matches: 451  
 Percent Similarity: 46.20% Conservative: 266  
 Best Local Similarity: 29.06% Mismatches: 551  
 Query Match: 19.60% Indels: 284  
 DB: 13 Gaps: 59  
 US-10-020-095-4 (1-1428) x US-10-292-081A-3 (1-4577)  
 QY 12 LeuLeuCysValCysThrAlaAlaLeuAlaValAlaProGlyProArgPheLeuValThr 31  
 DB 86 CTCTGGTCTCTCTCCACAGAGCCCTCAGCTCTGGAAGAAACCGCAGATATATGTTCTG 145  
 QY 32 AlaProGlyIleIleArgProGlyValAsnValThrIleGlyValGlnLeuGlnHis 51  
 DB 146 GTCCCTCCCTCTCTCACACT---GAGACCAGTGAAGAGGCTGTGCTCTTGAAGTAC 202  
 QY 52 CysProSerGlnValThrValLysAlaGlnLeuLeuLysThrAlaSerAsnLeuThrVal 71  
 DB 203 CTGAATGACAGACGTACTGTAAGTCTCTTGGAGTCTGTCAAGGAAACAGAGACCTC 262  
 QY 72 ---SerValLeuGlnAlaGlnGly---ValPheGlnLysGlySerPheLysThrLeuThr 89  
 DB 263 TTCACTGACCTGAGAGCGGAGAAATGACGTACCTCACTGTGTGCTTC----- 310  
 QY 90 LeuProSerLeuProLeuAsnSerAlaAspGlu-----IleTrpGlnLeuArgVal 106  
 DB 311 ---GCTGTCCAAAGCTTTCATTCATATGAGAGGTATGTTCTCTCAAGTGAAG 364  
 QY 107 ThrGlyArgThrGlnAspGlnLeuLeuPheSerAsnSerThrArgLysSerPheGluThr 126  
 DB 365 AAGAGCAACCAAGAA---TTTAAAGACCGGACACACAGTATGTTTAAAGAAC 415  
 QY 127 LysArgIleSerValPheIleGlnThrAspLysAlaLeuTrpLysProLysGlnGluVal 146  
 DB 416 GAGGACAGTGTGCTTGTGTCCAGACAGACAAATCAATCTCAAAACCAAGGACAGAGTG 475

QY 147 LysPheArgIleValThrIleuPheSerAspPheLysProTyrLysThrSerLeuAsnIle 166  
 Db 476 AAATTTCCTGTCCTCATGGATGAAACATTTCACCCCTGAATGATTCATTCACCTA 535  
 QY 167 Leu--IleLysAspProLysSerAsnLeuIleGlnIntProLysSerGlnIleSerAsp 185  
 Db 536 GTATTCATTTCAGATCCCAAGAAATGCGATCGCAATGCGAGAGTTTCAGTTAGAG 595  
 QY 186 LeuGlyValIleSerLysThrPheGlnIleuSerSerHisProIleLeuGlyAspTyrSer 205  
 Db 596 GGTGGCTTCAGCAATTTCTTTCTTCCTCATACAGAGCCCTTCAGAGGCTCCACAG 655  
 QY 206 IleGlnIleGlnValAsnAsp-----GlnThrTyrTyrGlnSerPheGlnValSerGlu 223  
 Db 656 GTGCGTCATCAAGAAATTCAGTCGGAAGAGACAGACAC-----CCTTACCGCTGAGAGAA 712  
 QY 224 TyrValIleuProLysPheGlnValThrLeuGlnThrProLeuTyrCysSerMetAsnSer 243  
 Db 713 TTTGTTCTTCCCAAGTTGAAGTACAGTAACAGTGCACAAAGATATACATCATCTTGGAA 772  
 QY 244 LysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProValIleGlyAsp 263  
 Db 773 GAAGAGATGAATGTATCAGGTGTGGCTTATACATATGAGAAAGCTGTCCCTGGACAT 832  
 QY 264 ValThrIleuThrPheLeuProLeu----- 271  
 Db 833 GTGACTGTGACATTTGGCAAGAAATGATAGTACGCTTCCGACTGCCAGCGTGAAGATTCA 892  
 QY 272 -----SerPheTyrGlyLys 276  
 Db 893 CAGGCTTCTGTGAATAATTCAGTGCAGCTAAACAGCCATGGCTGCTTCATACAGCA 952  
 QY 277 LysLysAsnIleThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAsp 296  
 Db 953 GTAAAA-----ACCAAGCTCTTCAGCTGAAGAG-----AAGAG 988  
 QY 297 GluIleMetLysAsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSer 316  
 Db 989 TATGAAATGAAA-----CTTCACACTGAGAGCCCAAGATCAA 1024  
 QY 317 SerProGlyProValGluIleLeuThr-----ThrValThrGlnSerVal 331  
 Db 1025 GAAGAGAGAACAGTGTGGAATTGACTCGAAGAGCAGCTCCAGTGAATATCAAGAAACATA 1084  
 QY 332 ThrGlyIleSerAspAsnValSerThrAsnValPhePheLysGlnHisAspTyrIleIle 351  
 Db 1085 ACCAAACCTCA-----TTTGTAAAGTGACTCACACTTCGACAGGGA-----ATT 1132  
 QY 352 GluPhePheAspTyrThrThrValLeuLysProSerIleAsnPheThrAlaThrValLys 371  
 Db 1133 CCCTTCTTTGGG-----CAG 1147  
 QY 372 ValThrArgAlaAspGlyAsnGlnLeuThrLeuGlnGluArgAsnAsnValIle 391  
 Db 1148 GTGGCCCTAGTATGATGAGAAAGCGTCCCTATACCA-----AATAAAGTCTATA 1195  
 QY 392 ThrValThrGlnAspAsnTyrThrIuTyrTyrSerGlySerAsnSerGlyAsnGlnLys 411  
 Db 1196 TTTCATC--AGAGGAATGAAGCAAACTATTAC-----TCCAAATGCTTACCAAGGATAG 1246  
 QY 412 MetGluAlaValGln--LysIleAsnTyrThr-----ValProGlnSerGlyThrPhe 428  
 Db 1247 CATGCCCTTGTATCACTTCTTATTCACACACCAACAGTTATGGGACTCTCTTATCTTT 1306  
 QY 429 LysIleGluPhePro-----IleLeuGluAspSer 438  
 Db 1307 AGGGTCATTTACAGAGATCGTATGCTTCGTTACGGCTACAGTGGGTCTCAAGAAACAC 1366  
 QY 439 SerGluLeuGlnLeuLysAlaTyrPheLeuGlySerLysSerMetAlaValHisSer 458  
 Db 1367 GAAGGAGCACATCACTGCTTAT----- 1390

QY 459 LeuPheLysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLys 478  
 Db 1391 CTTGTGTTCTCCCAACAGAGAGCTTTGTCCACTGAGCCCATGTCTCATGAACCTACCC 1450  
 QY 479 ValGlySer-----ProPheGlnLeuValValSerGlyAsnLys-----Arg 492  
 Db 1451 TGTGGCCATCTCAGACAGTCCAGGACCATTAATTTGATGATGAGGACCCCTGCTGGG 1510  
 QY 493 LeuLysGluLeuSer-----TyrMetValIleSerArgGlyGlnLeuValAlaValGly 510  
 Db 1511 CTGAAGAGCTCTCTTTTATCTATATGACCAAGGAGGACATGTCGCAACTGGG 1570  
 QY 511 -----LysGlnAsnSerThr-----MetPheSerLeuThr 520  
 Db 1571 ACTCATGAGCTGCTGTGAAGAGAGAGACATGAAGGCCATTTTTCATCTCATCTCCCT 1630  
 QY 521 ProGluAsnSerThrProLysValAlaCysValIleValAlaTyrTyrIleGluAspArgGly 540  
 Db 1631 GTGAAGTCAGACATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1690  
 QY 541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560  
 Db 1691 GACGTGATTTGGGATTTCTGCMAAATATGATGTTGAAATTTGCTGCGCAACAGTGGAT 1750  
 QY 561 LeuTyrTyrSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580  
 Db 1751 TTGAGCTTACGCCCATCAACAAAGTCTCCAGCTTCACACCCCACTGCGAGTCAACGCG 1810  
 QY 581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600  
 Db 1811 GCT---CCTCAGTCCGCTCGCGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1867  
 QY 601 AlaserAsnAspIleThrMetGluAsnValValHisGluLeu---GluLeuTyrAsnThr 619  
 Db 1868 CCTGATCTGAGCTCTGCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1927  
 QY 620 GlyTyr-----TyrLeuGlyMet 625  
 Db 1928 GGCTTCCCTGGGCTTGAATGACACAGACGATGAAGATCATCATCTCATTAATGTG 1987  
 QY 626 PheMetAsnSer-----PheAla 631  
 Db 1988 TATATTATGTAATCATCATATACTCCAGTATCAAGTACAAATGAAAGATATGTATACG 2047  
 QY 632 ValPheGlnGluCysGlyLeuTyrValIleuThrAspAlaAsnLeuThrLysAspTyrIle 651  
 Db 2048 TTCTTAGAGACATGGGCTTAAAGCATTCACCACTCAAAAGATTGCTTAACCAAAATG 2107  
 QY 652 AspGlyValTyrAspAsnAlaGlu-----Tyr 660  
 Db 2108 TGTCCACAGCTTCAACAGTATGAATGCATGACCTGAAGGTCTAGCTGATGTTTAT 2167  
 QY 661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680  
 Db 2168 GAGTCAGATTAATGGAGAGAGGCGCATGACGCGCTGATGATGTGAAGG----- 2218  
 QY 681 GlySerSerProHis-----ValArgLysHisPheProGluThrTyrIleTyrLeu 697  
 Db 2219 -----CCTCACAGGAGACCGGCTTGAAGAGTCTCTGAGCATGATCTGGAGT 2269  
 QY 698 AspThrAsnMetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIle 717  
 Db 2270 TTGGTGTGTGAATCAAGAGAGGAGGCTGAGAGTGAATTAACAGTCCCTGACACCATC 2329  
 QY 718 ThrSerTyrAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThr 737  
 Db 2330 ACCGAGTGAAGGAGGAGGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2389  
 QY 738 ThrProValGluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSer 757  
 Db 2390 ACT---GCTCTCTCCAGGCTTCCAGCTTCTTGTGAAGCTTCAAGAGCTTACTCT 2446  
 QY 758 ValIleArgGlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAla 777

Db 2447 GTATTCGTGGAGGCTTCACACTCAAGCCAGCTCTCAAACTTCCCAATGC 2506  
 Qy 778 ThrGluValIleIleGluYserAspIleLeuMetThrSer 797  
 Db 2507 ATCCGGGTGAGTGTGACGTGGAAGCTCTCCGCCCTTCTCTGCTGCCAGTGGAGAG 2566  
 Qy 798 Glu-----IleAsnAlaThrGlyHisGlnGlnThrLeuValProSer 812  
 Db 2567 GAACAAGCCGCTCACTGCATCTGTGCMAACGGCGGCA----- 2605  
 Qy 813 GluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGlyIleProIle 832  
 Db 2606 -----ACTGTCTCTGGGCACTAACCCCAAGTCATTAGGAATGTGAATTC 2653  
 Qy 833 ThrValThrAlaLeuSer-----ProThrAla 841  
 Db 2654 ACTGTAGCCCAAGGCACTAGAGCTCAAGAGCTGTGGAGCTGAGGCTTCACTT 2713  
 Qy 842 -----SerAspAlaValThrGlnMetIleLeuValIleValIleGluGlyIle 856  
 Db 2714 CTGAACACGGAGAGAAAGACACAGTCATCAAGCTCTGTGGTTGAACCTGAAGACAT 2773  
 Qy 857 GluYserTyrSerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThr 876  
 Db 2774 GAGAAAGAAACAACTTCACTCTTGTCTCACTGAGTGTGAGGTTCT----- 2827  
 Qy 877 LeuValThrLeuSerPheSerPheProProAsnThrValThrGlySerGluValGln 896  
 Db 2828 ---GAAGAATTTATCCCTGAACCTGCCAACAATGTGTAGAGAAATCTGCCAGCTTCT 2884  
 Qy 897 IleThrAlaIleGlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArg 916  
 Db 2885 GTCTCAGTTTGGAGACATATTAGGCTCTGCCATGCAAAACACAAATCTTCTCCAG 2944  
 Qy 917 MetProTyrGlyCysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeu 936  
 Db 2945 ATGCCCTATGGCTGTGAGAGAGCAAAATGTGCTCTTGTCTCAATCTATGATCAG 3004  
 Qy 937 AspTyrLeuThrIleValIleValIleValIleValIleValIleValIleValIleVal 956  
 Db 3005 GATTATCTAATTAATAACACAGCAGCTTACTCCAGAGCTCAAGTCCAGGCAATGTGCTAT 3064  
 Qy 957 MetArgGlnGlyTyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAla 976  
 Db 3065 CTCAACACTGTGTTACAGAGACAGCTTGAATCAAAACATATATGCTCTCAACAGACC 3124  
 Qy 977 PheGly---AsnTyrAspProSer---GlySerThrTyrLeuSerAlaPheValLeuArg 994  
 Db 3125 TTGGGGAGGATATGAGCAAGAACAGGCAACCTGGCTCAAGCTTGTGTGAG 3184  
 Qy 995 CysPheLeuGluAlaAspProTyrIleAspIleAspGlnAsnValLeuHisIleArgTyr 1014  
 Db 3185 ACTTTGGCCCAAGCTCAGAGCTCACTTCATGTAGTAAGACACATCCCAAGCCCTC 3244  
 Qy 1015 ThrTyrLeuIleGlyHisGlnYserAsnGlyGluPheTyrAspProGlyArgValIle 1034  
 Db 3245 ATATGCTCTCCAGAGGCAAGAGACATGCTGTTTCAAGAGCTCTGGGTACAGCTC 3304  
 Qy 1035 HisSerGluLeuGlnGlyIleAsnYserProValThrLeuThrAlaTyrIleValThr 1054  
 Db 3305 AACAAATGCCATTAAGAGAGAGTAAGTAAGTAGACCTCTCCGCTATATACACATC 3364  
 Qy 1055 SerLeuLeuGlyTyrArgIleTyrGlnProAsnIleAspValGlnLeuSerIleHisPhe 1074  
 Db 3365 GCCCTTCTGAGATCTCTCAAGTCACTCACTCCTGTGTGCGCATGCTGTTTTCG 3424  
 Qy 1075 LeuGluSer-----GluPheSerArgGlyIleSerAspAsnTyrThr 1088  
 Db 3425 CTGAGTCAAGCTCGGAAGACAGACAAAGAGGGGACATGAC---AGCCATGATATAC 3481  
 Qy 1089 LeuAlaLeuIleThrTyrAlaLeuSerSerValGlySer---ProValAlaValGluAla 1107

Db 3482 AAAGCATCTGCGCTTATGCTTTTGGCTTGGAGGTAACAGACAGAGAGAGTA 3541  
 Qy 1108 LeuAsnMetLeuThrTyrArgIleGlnGlnGlyIleMetGlnPheThr----- 1124  
 Db 3542 CTCAAGTCACTTAATAGAGAACTGTGAAGAAAGACAACTGTCTCATTTGGAGCGCCT 3601  
 Qy 1125 ValSerSerGluSerTyrLeuSerAspSerTyrGlnProArg-----SerLeuAspIle 1142  
 Db 3602 CAGAAACCAAGGACACAGATGGGGCATTTTATGAAACCCAGAGCTCCCTGTGTAGAGTG 3661  
 Qy 1143 GluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGln-----ThrSer 1159  
 Db 3662 GAGATGACATCTTAATGTGCTCTCGCT---TATCTCACGCGCCAGCAGACCCCACTCG 3718  
 Qy 1160 GluGlyIlePro-----IleMetArgTyrPheSerArgGlnArgAsnSerLeu 1175  
 Db 3719 GAGAGCTTAACTCTGCAACCAATCGAAGTGTATACAGAGAGAGATGCCAG 3778  
 Qy 1176 GlyGlyPheAlaSerThrGlnAspThrThrValAlaAlaLeuValAlaLeuSerGluPheAla 1195  
 Db 3779 GCGGTTTCTCTCCACCCAGAGACAGTGTGTGCTCTCATGCTCTGTCCAAATATGGA 3838  
 Qy 1196 AlaLeuMetAsnThrGluArgThrAsnIleGlnValThrValThrGlyProSerSerPro 1215  
 Db 3839 GCCGCCACATTTTACC---AGGACTGGGAAGCTGCAAGTGTACTATCCAGTCTTACAGGG 3895  
 Qy 1216 SerProLeuAlaValValGln----- 1222  
 Db 3896 ACATTTTCCAGCAAAATTCCAAGTGTGACAAACAATGCCCTGTACTGACAGAGTCTCA 3955  
 Qy 1223 -----ProMetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCys 1238  
 Db 3956 TTGCCAGAGCTCTCGGGATATACAGCAGTAAGAGACAGAGAGAGATGTCTTACCTC 4015  
 Qy 1239 GlnLeuAsnValValTyrAsnValIleValIleValIleValIleValIleValIleVal 1258  
 Db 4016 CAGACCTCTTCAATAATACAAATAT-----CTC 4042  
 Qy 1259 GlnAsnGlnAlaAlaPheAspLeuAspValAlaValIleValIleValIleValIleVal 1275  
 Db 4043 CCAGAAAGAGAGAGTCTCCCTTGTCTTGAAGTGCACACTGTGCTCAAACTGTGTAT 4102  
 Qy 1276 AspLeuAsn---HisValAspLeuAsnValCysThrSerPheSerGlyProGly----- 1292  
 Db 4103 GAACCCAAAGCCCAACAGCAGCTTCCAATCTCCCTAAGTGTATGATACAGAGGCCGC 4162  
 Qy 1293 ---ArgSerGlyMetAlaLeuMetGluValAsnLeuLeuSerGlyPheMetValProSer 1311  
 Db 4163 TCTGCCTCCAAATGCGCATGCTGTATGTGAAGATGTCTGTGCTTCACTTCCCTGAAAG 4222  
 Qy 1312 GluAlaIleSerLeu-----SerGluThrValIleValIleValIleValIleValIleVal 1328  
 Db 4223 CCAGACGTAAATCTTGAAGATCTTAACCATGTGAGCGCGGACAGAAATCGACAGCAAC 4282  
 Qy 1329 LysLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCysValAsnIleProAla 1348  
 Db 4283 CATGTCTTCAATTAATCTGATTAAGTGTCAAAATCGACACTGAGCTGTGTTCTTCAAGCTT 4342  
 Qy 1349 ValArgAsnPheValSerAsnThrGlnAspAlaSerValSerIleValAspTyrTyr 1368  
 Db 4343 CTGCAAGATGTCCAGTAAGATCTCAAAACAGCATGTAGAAAGTATATGATTACTATC 4402  
 Qy 1369 GluProArgArgGlnAlaValArgSerTyrAsnSer 1380  
 Db 4403 GAGACGATGAGTTTGCATTCGCTGAGTAATGCT 4438

RESULT 12  
 US-10-292-081A-5  
 ; Sequence 5, Application US/10292081A  
 ; Publication No. US20030162202A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kenneth David Becker  
 ; APPLICANT: Gonnul Vellicelebi

APPLICANT: Xin Wang  
 APPLICANT: Randolph E. Tanzi  
 APPLICANT: Lars Bertram  
 APPLICANT: Aleister J. Saunders  
 TITLE OF INVENTION: SINGLE NUCLEOTIDE POLYMORPHISMS AND MUTATIONS ON ALPHA-2-MACROGLOBULIN  
 FILE REFERENCE: 37481-3323  
 CURRENT APPLICATION NUMBER: US/10/292, 081A  
 PRIOR FILING DATE: 2002-11-08  
 PRIOR APPLICATION NUMBER: 60/337434  
 NUMBER OF SEQ ID NOS: 15  
 SOFTWARE: FASTSEQ for Windows Version 4.0  
 SEQ ID NO: 5  
 LENGTH: 4577  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 US-10-292-081A-5

Alignment Scores:  
 Pred. No.: 3,56e-133 Length: 4577  
 Score: 1440.00 Matches: 451  
 Percent Similarity: 46.20% Conservative: 266  
 Best Local Similarity: 29.06% Mismatches: 551  
 Query Match: 19.60% Indels: 284  
 Gaps: 59

US-10-020-095-4 (1-1428) x US-10-292-081A-5 (1-4577)

QY 12 LeuLeuCyValAlCyThrAlaAlaLeuAlaValAlaProGlyProAlaGlyPheLeuValThr 31  
 DB 86 CTCCTGTCCTCCGCCACAGAGCGCTCAGTCTCTGAAAACGGCATATATGTTCTG 145  
 QY 32 AlaProGlyIleIleArgProGlyGlyValAsnValThrIleGlyValIleLeuGluHis 51  
 DB 146 GTCCCTCCCTCCCTCCACACT--GAGACCACTAGAGAGCGCTGCTCTTCTAGCTAC 202  
 QY 52 CysProSerGlnValThrValAlaGluLeuLeuValThrAlaSerAsnLeuThrVal 71  
 DB 203 CTGAATGAGACAGCTGATGATGCTTCTTGAAGTCTGACAGGAGAAACAGAGAGCTC 262  
 QY 72 ---SerValLeuGluAlaGluGly---ValPheGluValSerPheLeuThr 89  
 DB 263 TTCACTACCTGAGGCGGAGAAATGACGTACTCCACTGTGTGCGCTTC----- 310  
 QY 90 LeuProSerLeuProLeuAsnSerAlaAspGlu-----IleTyrgIleLeuArgVal 106  
 DB 311 -----GCTGTCCCAAGTCTTCAATCAATGAGAGGATGTTCTTCACTGCTCAAGTG 364  
 QY 107 ThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThrArgLeuSerPheGluThr 126  
 DB 365 AAAGCAACCAACCAAGAA-----TTTAAAGAGGAGACCAACATGATGTTAAGAAC 415  
 QY 127 LysArgIleSerValPheIleGlnThrAspIleValLeuTyrgIleProLysGlnGluVal 146  
 DB 416 GAGACAGCTGTGCTTGTTCACAGACAAATCAATCAAAACCAAGGACAGACAGCTG 475  
 QY 147 LysPheArgIleValIleThrLeuPheSerAspPheLeuProTyrgIleThrSerLeuAsnIle 166  
 DB 476 AAATTTGTGTGTCTCATGATGATAAACTTCACTCCCTGAATGATGATTCACATA 535  
 QY 167 Leu---IleLysAspProLysSerAsnLeuIleGlnIleThrLeuSerGlnIleSerAsp 185  
 DB 536 GTATACATTTCAGATCCCAAGAAATGCGCATGCGACATGGCATGTTCCAGTTAGAG 595  
 QY 186 LeuGlyValIleSerLysThrPheGluLeuSerSerHisProIleLeuGlyAspTyrSer 205  
 DB 596 GGTGGCTCAAGCAATTTCTTTCCCTCTCATCAAGCCCTTCCAGGCGCTCTTACAG 655  
 QY 206 IleGlnValGlnValAsnAsp-----GlnThrTyrgIleThrIlePheGlnValSerGlu 223  
 DB 656 GTGTGTGACGAAGAAATCAGGTGAGAGACAGACAC--CCTTTCACCGGTGAGGAA 712  
 QY 224 TyrValLeuProLysPheGluValIleThrLeuGlnThrProLeuTyrgCysSerMetAsnSer 243

DB 713 TTGTTCTTCCCAAGTTGAGATGACAAAGTAAACAGTCCCAAGATATCACTTCGAA 772  
 QY 244 LysHisIleLeuAsnGlyThrIleThrAlaLysTyrgIleValProValLysGlyAsp 263  
 DB 773 GAGAGATGATATTCAGATGATGTTGGCTTATACATATGGAAGGCTGTCTCCGACAT 832  
 QY 264 ValThrLeuThrPheLeuProLeu----- 271  
 DB 833 GTGACTGTGACATTTGCAAGAAATATAGACCTTCCGACTGCCAGGTGAAGATTCA 892  
 QY 272 -----SerPheTyrGlyLys 276  
 DB 893 CAGCTTTCTGTGAGAAATTCAGTGCAGCTTAACAGCCATGCTGCTTATTCAGCA 952  
 QY 277 LysValAsnIleThrIleThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAsp 296  
 DB 953 GTAAAA--ACCAAGCTCTCCAGCTGAGACAG--AGGAG 988  
 QY 297 GlnGluMetLysAsnValMetAspSerSerAsnGlyLeuSerGluTyrgIleAspLeuSer 316  
 DB 989 TATGAATATGAA-----CTTCACACTGAGGCCAGATCCAA 1024  
 QY 317 SerProGlyProValGluIleLeuThr-----ThrValThrGluSerVal 331  
 DB 1025 GAAGAAGCAACAGTGTGAGATGACTGGAAGGCACTCCAGTGAATTCACAGAACCATTA 1084  
 QY 332 ThrGlyIleSerArgAsnValSerThrAsnValPhePheGluHisAspTyrgIleIle 351  
 DB 1085 ACCAAATCTCA--TTTGTGAAAGTGCACACTCATTTGAGAGGA-----ATT 1132  
 QY 352 GluPhePheAspTyrgIleThrValLeuLysProSerLeuAsnPheThrAlaThrValLys 371  
 DB 1133 CCCTTCTTGGG-----CAG 1147  
 QY 372 ValThrArgAlaAspGlyAsnGlnIleThrLeuGluGluArgArgAsnValValIle 391  
 DB 1148 GTCCGCTGATGATGAGAAAGCGTCCCTATACCA-----ATAAAGTCATA 1195  
 QY 392 ThrValThrGlnArgAsnTyrgIleTyrgIleTyrgIleSerAsnSerGlyAsnGlnLys 411  
 DB 1196 TTCAATC--AGAGAAATGAGCAAACTATATC-----TCAATCTACACCGATGAG 1246  
 QY 412 MetGluAlaValGln--LysIleAsnTyrgIle-----ValProGlnSerGlyThrPhe 428  
 DB 1247 CARGGCTGTATCAGTCTCTATCAACACCAACAGTATAGGTACTCTTACTTACGTT 1306  
 QY 429 LysIleGluPhePro-----IleLeuGluAspSer 438  
 DB 1307 AGGTCATATACAGAGATGATGCTGCTTACGCTTACAGACAGTGGGTGCAGAAAGAAC 1366  
 QY 439 SerGlnLeuGlnLeuLysAlaTyrgIleLeuGlySerLysSerSerMetAlaValHisSer 458  
 DB 1367 GAAGAGCACATCACCTCTAT----- 1390  
 QY 459 LeuPheLysSerProSerLysThrTyrgIleGlnLeuLysThrArgAspGluAsnIleLys 478  
 DB 1391 CTGTGTCTTCCCAAGCAAGAGCTTGTCCACTTGAACCCATGTCTGATGAACTACCC 1450  
 QY 479 ValGlySer-----ProPheGluLeuValIleSerGlyValLys-----Arg 492  
 DB 1451 TGTGGCCATATCTACAGACGTCCAGACATTAATCTGAATGAGAGCACCTGTGCGAG 1510  
 QY 493 LeuLysGluLeuSer-----TyrMetValIleSerArgGlyGlnLeuValAlaValGly 510  
 DB 1511 CTGAAGAACTCTCTTTATATATGATTAATGCAAGAGAGGAGCTTGTCCAGACTGGG 1570  
 QY 511 -----LysGlnAsnSerThr-----MetPheSerLeuThr 520  
 DB 1571 ACTCATGACGTGTTGTGAAGACGAAGACATGAAGGCAATTTTCCATCTCAATCCCT 1630  
 QY 521 ProGluAsnSerTyrThrProLysAlaCysValIleValTyrgIleGluAspAspGly 540  
 DB 540 ----- 540



Db 1631 GTGAGTACAGACATTGCTCTGCTCGGTGGTTCATCTATGCTGTTTACACCGGG 1690  
 Qy 541 GGuillelleserAspValleuylsileProValGlnleuValPheylsasnlylleyls 560  
 Db 1691 GACGTGATTTGGGGATTTGCGCAAAATATATGTTGAAATTTCTGGCCACAGAGTGAT 1750  
 Qy 561 LeuTYrTSPSerLysVallysalagluProSerGluLysValSerleuArglleSerVal 580  
 Db 1751 TTGAGCTTCAGCCCATCAACAAAGTCTCCAGCCTCAACGCCCATCTGGAGTCAAGG 1810  
 Qy 581 ThrGlnProAspSerlleValGlylleValAlaValAspLysSerValAsnleuMetAsn 600  
 Db 1811 GCT---CCTCAGTCCGCTGCGCCCTCCGCTGCTGGACCAAGCGTCTGTCTGAG 1867  
 Qy 601 AlaserAsnAspIleTherMetGluAsnValAlaHisgluLeu---GluLeuTYrAsnThr 619  
 Db 1868 CCGTATGCTGACGCTCTCGGCGCTCTGCTTACAACTGCTTACAGAAAGAGCCTCACT 1927  
 Qy 620 GlyTYr-----TyrleuGlyMet 625  
 Db 1928 GGCTTCCCTGGGCTTTGAATGACAGACGATGAAGACTGCATCATGTCATATGTC 1987  
 Qy 626 PheMetAsnSer-----PheAla 631  
 Db 1988 TATATTAAATGAGATCATATATCTCCAGTATCAAGTACAAATGAAGATATGTACAGC 2047  
 Qy 632 ValPheGlnGluCyseGlyLeuTYrValleuThrAspAlaAsnleuThrLysAspTYrIle 651  
 Db 2048 TTCTTAGAGACATGGGCTTTAAAGGCAATTCACCAACTCAAGATTCGTAACCCAAATG 2107  
 Qy 652 AspGlyValTYrAspAsnAlaglu-----Tyr 660  
 Db 2108 TGTCCACAGCTTCAACAGTATGTAATGATGATGACCTGAAGGCTACGCTAGCTTTTAT 2167  
 Qy 661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerleu 680  
 Db 2168 GAGTCAGATGATGATGAGAGAGGCGCATGACCGCTGCTGATGTTGAAG----- 2218  
 Qy 681 GlySerSerProHis-----ValArgLysHisPheProGlnThrTrioIleThreleu 697  
 Db 2219 -----CCTCACACGAGACCGGTACGAAAGTACTCTCCGAAACATGATGCTGGAT 2269  
 Qy 698 AspThrAsnMetGlyTYrArgIleTYrGlnGluPheGluValThrValProAspSerlle 717  
 Db 2270 TTGGTGTGTAACTCAGACAGGGGCTGCTGAGTAAAGATACAGTACCTCTACACATC 2329  
 Qy 718 ThrSerTPValAlaThrGlyPheValIleSerGluAspLeuGlyLeuThrThr 737  
 Db 2330 ACCGAGTGAGAGGAGGCGGCTTCTGCTGCTGAAGATGCTGACTTGTATCTCTCC 2389  
 Qy 738 ThrProValGluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnleuProTYrSer 757  
 Db 2390 ACT---GCCCTCTCCGAGCCTTCCAGCCCTTTTGTGAGCTTACATCAATGCTTACTCT 2446  
 Qy 758 ValIleArgGlyGluGluPheAlaLeuGluIleThrIlePheAsnTYrleuLysAspAla 777  
 Db 2447 GTGATTCGTGAGAGGCTTTCACACTCAAGGCCACGCTCTTAACCTTCCCAATGTC 2506  
 Qy 778 ThrGluValLysValIleleGluLysSerAspLysPheAspIleleuMetThrSerSer 797  
 Db 2507 ATCCGGGTGAGTGTGAGAGGCTCTCCGCTCTGCTGCTGCTCCAGTGAAG 2566  
 Qy 798 Glu-----IleAsnAlaThrGlyHisGlnGlnThrleuLeuValProSer 812  
 Db 2567 GAACAAGCCCTCACTGCACTGTGTGCAAAAGCGGCGCA----- 2605  
 Qy 813 GluAspGlyAlaThrValLeuPheProIleArgProThrHisIleGlyGluIleProIle 832  
 Db 2606 -----ACTGTTCCTGGGAGTAAACCCCAAGTCAATTAGAAATGTGAATTC 2653  
 Qy 833 ThrValThrAlaLeuSer-----ProThrAla 841  
 Db 2654 ACTGTGAGGCGAGAGGCACTAGAGTCTCAAGAGCTGTGTGGAGCTGAGGTGCTTCAAT 2713

Qy 842 -----SerAspAlaValThrGlnMetIleLeuVallysalagluGlyIle 856  
 Db 2714 CCGTAACAGGAGAAAGAACACAGTCAATCAAGCTCTGTTGTTGTAACCGAAGACTA 2773  
 Qy 857 GluLysSerTYrSerGlnSerIleLeuLysAspLeuThrAspAsnArgleuGlnSerThr 876  
 Db 2774 GAGAAAGAAACAAATTCATCACTCTTGTCTCATCAGGTGTGAGGTTCT----- 2827  
 Qy 877 LeuLysThrleuSerPheSerPheProProAsnThrValThrGlySerGluArgValGln 896  
 Db 2828 ---GAGAAATTAATCTGAAATCTGCACCAATGTGTAGAGATCTGCCGAGCTTCT 2884  
 Qy 897 IleThrAlaIleGlyAspValleuGlyProSerIleAsnGlyleuAlaSerleuIleArg 916  
 Db 2885 GTTCACAGTTTGGAGACATATTAGCTCTGCTCCATGCAAAAACACAAAATCTTCTCAG 2944  
 Qy 917 MetProTYrGlyCyseGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTYrIleLeu 936  
 Db 2945 ATGCCCTATGCTGTGGAGACAGATATAGTCTCTTGTCTCTTACATCTATGATCTG 3004  
 Qy 937 AspTYrleuThrLysLysGlnleuThrAspAsnleuLysGlyValAlaLeuSerPhe 956  
 Db 3005 GATTATCTAAATGAAGACAGACAGCTTACAGAGGTCAAGTCCAAAGGCCATGGCTAT 3064  
 Qy 957 MetArgGlnGlyTYrGlnArgGluLeuLeuTYrGlnArgLysAspLysSerPheSerAla 976  
 Db 3065 CTCACACTGTTTCCAGACAGCTTACACTTACAAACCTATGATGGCTCTTACAGACC 3124  
 Qy 977 PheGly---AsnTYrAspProSer---GlySerThrPheLeuSerAlaPheValleuArg 994  
 Db 3125 TTGGGAGAGCATATGGAGAGAACAGGCAACCTGGCTCACAGCCTTTGTTCTGAG 3184  
 Qy 995 CysPheleuGluAlaAspProTYrIleAspIleAspGlnAsnValleuHisArgThrTYr 1014  
 Db 3185 ACTTTTGCCCAAGCTCGAGCCTTACATCTTCATGATGAAGACACATTAACCAAGCCTC 3244  
 Qy 1015 ThrTYrleuLysGlyHisGlnLysSerAsnGlyGluPheThrAspProGlyArgValIle 1034  
 Db 3245 ATATGGCTCTCCAGAGGCAAGAAAGCAATGGCTGTTTACAGAGCTTGGGTCTGCTC 3304  
 Qy 1035 HisSerGluLeuGlnGlyGlyAsnLysSerProValThrleuThrAlaTYrIleValThr 1054  
 Db 3305 AACATGCCAATAAAGGAGAGAGTAAAGATGAATGACCTCTGCCCTATATACCATC 3364  
 Qy 1055 SerleuGluGlyTYrArgLysTYrGlnProAsnIleAspValGlnLysSerIleHisPhe 1074  
 Db 3365 GCCCTTCTGAGATTCCTCTCACAGTCACTACCCCTGTTGTCGCAATGCGCTTTTGC 3424  
 Qy 1075 LeuGluSer-----GluPheSerArgGlyIleSerAspAsnTYrThr 1088  
 Db 3425 CTGAGGTAGGCTGGAAGACAGCAAGAAAGGGACCATAGGC---ACCATGATATATACC 3481  
 Qy 1089 LeuAlaLeuIleThrTYrAlaLeuSerSerValGlySer---ProLysAlaLysGluAla 1107  
 Db 3482 AAAGCATGCTGGCTATGCTTTTGGCTCGGAGGTAAACAGCAAGAAAGAAAGTA 3541  
 Qy 1108 LeuAsnMetleuThrTPArgAlaGluGlnGluGlyLysMetGlnPheTP----- 1124  
 Db 3542 CTCAAGTCACTTATGAGAAAGCTGTGAAGAAACAACTGTGCAATTTGGAGCGCCCT 3601  
 Qy 1125 ValSerSerGluSerLysLysSerAspSerTProIleProArg-----SerleuAspIle 1142  
 Db 3602 CAGAAACCAAGGACAGCGGGGCAATTTTAAACCAACCCAGGCTCCCTGTGCTGAGTG 3661  
 Qy 1143 GluValAlaAlaTYrAlaLeuLysSerHisPheleuGlnPheGln-----ThSer 1159  
 Db 3662 GAGATGACATCTATGCTGCTCTGCT---TATCTCAAGGCCCAAGCAAGCCCACTCG 3718  
 Qy 1160 GluGlyIlePro-----IleMetArgTPLeuSerArgGlnArgAsnSerleu 1175  
 Db 3719 GAGAGCTGACCTCTGCAACCAACATCATGTGAAGATGATCAAGAGCAGAGATGCCAG 3778



QY 277 LyeLysAsnIleThrIysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAsp 296  
 Db 953 GTAAAA-----ACCAAGCTCTCCAGCTGAGAGG-----AAGAG 988  
 QY 297 GluGluMetLysAsnValMetLysSerSerAsnGlyLeuSerGluIyrLeuAspLeuSer 316  
 Db 989 TATGAATGANA-----CTTCACACTGAGGCCAGATCCAA 1024  
 QY 317 SerProGlyProValGluIleLeuThr-----ThrValThrGluSerVal 331  
 Db 1025 GAAGAAAGAAACAGTGTGGATTGACTGGAGAGCAGTCCAGTAATTCACAGAACATTA 1084  
 QY 332 ThrGlyIleSerArgAsnValSerThrAsnValPhePheLysGlnHisAspTyrIleIle 351  
 Db 1085 ACCAAACTCTCA---TTGTGAAGTGAAGTCACTTCGACAGGGA-----ATT 1132  
 QY 352 GluPhePheAspTyrThrThrValLeuLysProSerLeuAsnPheThrAlaThrValLys 371  
 Db 1133 CCCTTCCTGGG-----CAG 1147  
 QY 372 ValThrArgAlaAspGlyAsnGlnLeuThrLeuGluGluArgAsnAsnValAlaIle 391  
 Db 1148 GTCCGCTAGTAAATGGAAAGCGTCCCTATACCA-----ATAAAGTCATA 1195  
 QY 392 ThrValThrGluArgAsnTyrThrGluTyrTrpSerGlySerAsnSerGlyAsnGlnLys 411  
 Db 1196 TTCATC---AGAGAAATGAAGCAACTATTAC---TCCAATGCTACCAAGATGAG 1246  
 QY 412 MetGluAlaValGln---LysIleAsnTyrThr---ValProGlnSerGlyThrPhe 428  
 Db 1247 CATGCCCTGTATACATCTCTATACACACCAAGCTTAAGGTACTCTCTTACTGTT 1306  
 QY 429 LysIleGluPhePro-----LleLeuGluAspSer 438  
 Db 1307 AGGGCTAATTACAGATCTACTCCCTGTACGGCTACCAAGTGGGTGTCAGAGAACAC 1366  
 QY 439 SerGluLeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSer 458  
 Db 1367 GAAGAGGCATCATCACTGCTTAT----- 1390  
 QY 459 LeuPheLysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLys 478  
 Db 1391 CTGTGTCTCCCAAGCAAGAGACTTGTCCACCTTGAGCCCATGCTCATGAATACCC 1450  
 QY 479 ValGlySer-----ProPheGluLeuValValSerGlyAsnLys-----Arg 492  
 Db 1451 TGTGGCCATATCATGACAGTCCAGGACATTAATTCGTGAATGAGGACCACTGCGGG 1510  
 QY 493 LeuLysGluLeuSer-----TyrMetValValSerArgGlyGlnLeuValAlaValGly 510  
 Db 1511 CTGAAGAAGCTCTCTTTTATTTATCTGATATATGCAAGAGGAGGACATGTGCCAAGCTGG 1570  
 QY 511 -----LysGlnAsnSerThr-----MetPheSerLeuThr 520  
 Db 1571 ACTCATGAGCTGCTTGTGAAGCAGAAAGCATGAAAGGCGCATTTTTCATCTCAATCCCT 1630  
 QY 521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGluAspArgGly 540  
 Db 1631 GTGAAGTCAGACATGTGCTGTGCTGTGGTGTCTCATATGCTGTATTAACCTACCGGG 1690  
 QY 541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560  
 Db 1691 GAGCTGATTGGGATTCGTGCAAAATATGATGTTGAAAATGTGTGGCCAAAGAGTGAT 1750  
 QY 561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580  
 Db 1751 TTGAGCTTCAAGCCATCAAAAGTCTCCAGCTTCACAGCCCACTGCGAGTCACAGCG 1810  
 QY 581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600  
 Db 1811 GCT---CCTAGTCCGTGCGCCCTCGTGTGTCGAGCAAGCGTGTGCATGAGG 1867  
 QY 601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeu---GluLeuTyrAsnThr 619

Db 1868 CCTGATGCTAGACTCTCGCGCTCTGCTTATCAACCTGTCTACAGAAAGAGCTCTACT 1927  
 QY 620 GlyTyr-----TyrLeuLysMet 625  
 Db 1928 GGCCTCCCTGGGCTTTGAATGACCAAGACGATGAAGACTGCATCAATGCTATATGTC 1987  
 QY 626 PheMetAsnSer-----PheAla 631  
 Db 1988 TATATTAATGAAATCATATATCTCCAGATTCATACAAATGAAAGATATGTACAC 2047  
 QY 632 ValPheGlnGluCysGlyLeuTrpValLeuThrAspAlaAsnLeuThrLysAspTyrIle 651  
 Db 2048 TTCCTAGAGACATGAGGCTTAAAGCAATTCACCACTCAAGATTCGTAAACCAAAATG 2107  
 QY 652 AspGlyValTyrAspAsnAlaGlu-----Tyr 660  
 Db 2108 TGTCCACAGCTTCAACAGTATGAATGACATGAGCTGAAGCTTACGTATGTTTAT 2167  
 QY 661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680  
 Db 2168 GAGTCAGATGTATGGAAGAGGCAATGACAGCGCTGTGTCATGTGAAGAC----- 2218  
 QY 681 GlySerSerProHis-----ValArgLysHisPheProGluThrTyrIleThrPhe 697  
 Db 2219 -----CTCACAGGAGACCGTACGAAAGTACTTCCCTGAGACATGATCTGGAT 2269  
 QY 698 AspThrAsnMetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIle 717  
 Db 2270 TTGTGGTGTGTAACTCAGCAGGGGTGCTGAGTGAAGATTAACGTCCCTGACCAATC 2329  
 QY 718 ThrSerTrpValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGluThrThr 737  
 Db 2330 ACCGATGGAAGGCAAGGCGCTTGTGCTGTGTGAAGTCTGTGAGCTTACAAATGCTTCC 2389  
 QY 738 ThrProValGluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSer 757  
 Db 2390 ACT---GCCTCTCCAGAGCTTCCAGCGCTTCTTGTGAGACTTACAAATGCTTACTCT 2446  
 QY 758 ValIleArgGlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAla 777  
 Db 2447 GTGATTTGCGAGAGGCTTCCACTCACTCAAGGCAAGGCTCAATACCTTCCAAATGC 2506  
 QY 778 ThrGluValLysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSer 797  
 Db 2507 ATCCGAGTCAAGTGTACAGTGAACCTCTCCGCTTCTTGTGCTCCAGTGAAGAG 2566  
 QY 798 Glu-----IleAsnAlaThrGlyHisGlnGlnThrLeuValProSer 812  
 Db 2567 GAACAAGCGCTCAGCTGATCTGTGCAAAAGGCGGCA----- 2605  
 QY 813 GluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGlyGluIleProIle 832  
 Db 2606 -----ACTGTCTCGGCAAGTAAACCCCAAGTCAATTAGAAATGTGAATTC 2653  
 QY 833 ThrValThrAlaLeuSer-----ProThrAla 841  
 Db 2654 ACTGTAGCGGAGAGCACTAGAGTCTCAAGAGCTGTGTGGAGCTGAGGTGCTTCACTT 2713  
 QY 842 -----SerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIle 856  
 Db 2714 CCTGAACAGGAAGAAAGACACATCATCAAGCTCTGTGTGTAACCTGAAGACATA 2773  
 QY 857 GluLysSerTyrSerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThr 876  
 Db 2774 GAGAAGAAACAAACATTCAACTCTTACTTGTGCATAGGAGTGAAGATTCT----- 2827  
 QY 877 LeuLysThrLeuSerPheSerPheProProAsnThrValThrGlySerGluArgValGln 896  
 Db 2828 ---GAAGATTATCCCTGAACCTGCACCAATGTGTGAAGAAATGTGCCGAGACTTCT 2884  
 QY 897 IleThrAlaIleGlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArg 916

Db 2885 GTCTCAGTTTGGGAGACATATTAGGCTCTGCATCGAAAAACACAAATCTTCTCAG 2944  
 Qy MetProTyrGlyCySGlyGluGlnAsnMetIleasnPheAlaProAsnIleTyrIleu 936  
 Db 2945 ATGCCCTATGGCTGTGGAGACAGAAATATGGTCTCTTGGCTCTTAACATATATGTA 3004  
 Qy AspTyrLeuThrIlyValGlyGlnLeuThrAspAsnLeuIlyValAlaLeuSerPhe 956  
 Db 3005 GATTATCTTAATTAAGAAACACACAGCTTACTCCAGAGCTCAAGTCCAAAGGCATGGCAT 3064  
 Qy MetArgGlnIlyTyrGlnArgGlnLeuLeuTyrGlnArgIlyAspGlySerPheSerAla 976  
 Db 3065 CTCAACACTGGTATCAGAGACAGATTGAATCAACAACTATGATGGCTCTCTACAGACC 3124  
 Qy PheGly---AsnTyrAspProSer---GlySerThrTrpLeuSerAlaPheValLeuArg 994  
 Db 3125 TTTGGGAGCGCATATGGCAGGAAACCAAGCTTGGCTCTACAGCCTTTGTTCTAGAG 3184  
 Qy CysPheLeuGlnAlaAspProTyrIleAspIleAspGlnAsnValLeuIleAspThrTyr 1014  
 Db 3185 ACTTTGGCCCAAGCTCGAGCCTACATCTTCATCGATGAAGCACAATTACCCAAAGCCTC 3244  
 Qy ThrTrpLeuIlyGlyIleGlnIlySerAsnGlyGlnPheTrpAspProGlyArgValIle 1034  
 Db 3245 ATATGGCTCTCCAGAGCAGAAAGCAATAGCTGTTTCAGAGCCTCTGGTCACTGCTC 3304  
 Qy HisSerGlnLeuGlnIlyGlyAsnIlySerProValThrLeuThrAlaTyrIleValThr 1054  
 Db 3305 AACATGGCATTAAGAGAGAGTAGAAGATGAAGAGACCTCTCGCTTAATATCAACATC 3364  
 Qy SerLeuLeuGlyTyrArgIlyTyrGlnProAsnIleAspValGlnIlySerIleHisPhe 1074  
 Db 3365 GCCCTTGGAGATCTCTCAGAGTCACTCAGCCTGTTGTCGGAATGCCCTGTTTTCG 3424  
 Qy LeuGlnSer-----GlnPheSerArgGlyIleSerAspAsnTyrThr 1088  
 Db 3425 CTGAGTAGCCTCGAGAGACAGACACAAAGAGGAGCCATGGC---AGCCATGTAATATAC 3481  
 Qy LeuAlaLeuIleThrTyrAlaLeuSerSerValGlySer---ProIlyValIlyGlnAla 1107  
 Db 3482 AAAGCATCTGGCTATGCTTTTCCCTGGCAGGTAAACAGACAGAGAGAGAAAGTA 3541  
 Qy LeuAsnMetLeuThrTrpArgAlaGlnIlyGlnIlyMetGlnPheTrp----- 1124  
 Db 3542 CTCAAGTCACTTAATGAGAGAGCTGTGAAGAAAGCACTGTTCATGGAGAGCCCT 3601  
 Qy ValSerSerGlnSerIlyLeuSerAspSerTrpGlnProArg-----SerLeuAspIle 1142  
 Db 3602 CAGAAACCCAGAGCAGCAGTGGGCGATTTTAAAGAACCCAGGCTCCTCTGCTGAGAGTG 3661  
 Qy GluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGln-----ThrSer 1159  
 Db 3662 GAGATGACATCTTATGCTCTCTGCT---TATCTCAGGCGCCACAGCCCAACTCTG 3718  
 Qy GluGlyIlePro-----IleMetArgTrpLeuSerArgIlyAsnSerIleu 1175  
 Db 3719 GAGAGCTGACCTCGAACCAACATGTAAGTGAAGATCAAGAGAGCAAGCAATGCCAG 3778  
 Qy GlyIlyPheAlaSerThrGlnAspThrThrValAlaLeuIlyValAlaLeuSerGlnPheAla 1195  
 Db 3779 GGCAGTTCCTCTCCACCCAGAGACACAGTGGTGTCTCCATGCTCTCTCCAAATATATGCA 3838  
 Qy AlaLeuMetAsnThrGlnArgThrAsnIleGlnValThrValIlyPheSerSerPro 1215  
 Db 3839 GCCCGCACATTATAC---AGAGCTGGGAAGCTGCAAGGATCAATTCACAGTTCACAGGG 3895  
 Qy SerProLeuAlaValIlyGln----- 1222  
 Db 3896 ACATTTTCCAGCAATTCAGAGTGGACAAACAATGCGCTTACTGACAGAGTCTCA 3955  
 Qy PrometAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCys 1238  
 Db 3956 TTGCCAGAGCTGCTGGGGAATACAGCATGAAGTGAAGAGAAAGATGTGTCTACTTC 4015

Qy 1239 GlnLeuAsnValValTyrAsnValIlyValAspGlySerSerArgArgArgSerIle 1258  
 Db 4016 CAGAGCTCTCTGAAATATCAATATT-----CTC 4042  
 Qy 1259 GlnAsnGlnIlyAlaPheAspLeuAspValAlaValIly-----GluAsnIlyAsp 1275  
 Db 4043 CAGAAAGAGAAAGTTCCTCTTGGTGTAGAGTCAAGCTGTGCTCAAACTTGAT 4102  
 Qy AspLeuAsn---HisValAspLeuAsnValCysThrSerPheSerGlyProGly----- 1292  
 Db 4103 GAACCCAAAGCCACACACAGCTTCCAAATCTCCCTTAAGTGTCAAGTATACAGAGAGCCG 4162  
 Qy 1293 ---ArgSerGlyMetAlaLeuMetGluValAsnLeuSerGlyPheMetValProSer 1311  
 Db 4163 TGTGCTCCCAATGAGGAGATCGTTGATGTGAAGATGATGCTGTGCTTCAATTCCTGAG 4222  
 Qy 1312 GlnAlaIleSerLeu-----SerGluThrValIlyLeuIlyValGluTyrAspIleGly 1328  
 Db 4223 CCAACAGTGAATATGCTTGAAGATCTAACCATGTAGCCGAGACAGAAAGTCAAGCAAC 4282  
 Qy 1329 IlyLeuAsnLeuTyrIleAspSerValAsnGluThrGlnPheCysValAsnIleProAla 1348  
 Db 4283 CATGCTTGAATTAACCTTGTATAGGTGTCAAAATCAACACTGAGCTTGTCTTCAAGGT 4342  
 Qy 1349 ValArgAsnPheIlyValSerAsnThrGlnAspAlaSerValSerIleValAspTyrTyr 1368  
 Db 4343 CTGCAAGATGTCCAGTAAGAGATCTCAAAACGACCATGTGAAGTGTATGATTACTAC 4402  
 Qy 1369 GluProArgArgGlnAlaValArgSerTyrAsnSer 1380  
 Db 4403 GAGACGATGATGTTTGCATATCGCTGATCAATGCT 4438

## RESULT 14

US-10-240-965-178  
 ; Sequence 178, Application US/10240965  
 ; Publication No. US20030165924A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: INCYTE GENOMICS, INC.  
 ; APPLICANT: SHIFFMAN, DOV  
 ; APPLICANT: SOMOGYI, ROLAND  
 ; APPLICANT: LAMN, RICHARD M.  
 ; APPLICANT: SEILHAMER, JEFFREY J.  
 ; APPLICANT: PORTER, GORDON J.  
 ; APPLICANT: MIKITA, THOMAS  
 ; APPLICANT: TAL, JULIE  
 ; TITLE OR INVENTION: GENES EXPRESSED IN FOAM CELL DIFFERENTIATION  
 ; FILE REFERENCE: PA-0025 PCT  
 ; CURRENT APPLICATION NUMBER: US/10/240,965  
 ; CURRENT FILING DATE: 2002-10-04  
 ; PRIOR APPLICATION NUMBER: 60/195,106  
 ; PRIOR FILING DATE: 2000-04-05  
 ; NUMBER OF SEQ. ID NOS: 276  
 ; SOFTWARE: PERL Program  
 ; SEQ. ID NO 178  
 ; LENGTH: 4809  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; OTHER INFORMATION: Incyte ID No. US20030165924A1 1000033.6  
 US-10-240-965-178

## Alignment Scores:

Pred. No.: 3,87e-133 Length: 4809  
 Score: 1440.00 Matches: 451  
 Percent Similarity: 46.20% Conservative: 266  
 Best Local Similarity: 29.06% Mismatches: 551  
 Query Match: 13 Gaps: 284  
 DB: 59

US-10-020-095-4 (1-1428) x US-10-240-965-178 (1-4809)

QY 12 LeuLeuCysValCysThrAlaAlaValAlaProGlyProArgPheLeuValThr 31  
 DB 116 CTCTGTGCTCTCTGCCACAGACGCTCTCGAAGAACCGCATATATGTTCTCG 175  
 QY 32 AlaProGlyIleIleArgProGlyAlaValThrIleGlyValGluLeuGluHis 51  
 DB 176 GTCCCTCTCTCTCTCCACT--GAGACCACTGAGAGAGGCTGTCTCTCTGAGCTAC 232  
 QY 52 CysProSerGlnValThrValIleValIleLeuLeuLeuThrAlaSerAsnLeuThrVal 71  
 DB 233 CTGAATGAGACAGCTGACTGATGCTCTCTCTGAGCTCTGACGGGAAACAGAGCCCTC 292  
 QY 72 ---SerValLeuGluAlaGluGly--ValPheGluGlySerPheLeuThr 89  
 DB 293 TTTCAGCTGACCTGAGGCGGAGATGACGTAATCTCACTGTCGCTTC-- 340  
 QY 90 LeuProSerLeuProLeuAsnSerAlaAspGlu-----IleGlyGluLeuArgVal 106  
 DB 341 -----GCTGCCCAAGTCTTCATCCATGAGAGATGATCTTCCTCACTGTCCAAAGTG 394  
 QY 107 ThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThrArgLeuSerPheGluThr 126  
 DB 395 AAGAGCAACCAACCAAGAA-----TTTAAAGCGGACCAACAGTATGCTTAAAGAAC 445  
 QY 127 LysArgIleSerValPheIleGlnThrAspIleValAlaLeuThrIleProLysGlnGluVal 146  
 DB 446 GAGCAGACTGTGCTCTCTCTCCAGACAGCAATCAATCAATCAACAGGACAGAGTG 505  
 QY 147 LysPheArgIleValThrLeuPheSerAspPheLeuProTyrLysThrSerLeuAsnIle 166  
 DB 506 AATATTCGTGGTGTCTCCATGATGAAACCTTTCACCCCTGATGATGATGATTCACATA 565  
 QY 167 Leu---IleLysAspProLysSerAsnLeuIleGlnIleThrLeuSerGlnIleAsp 185  
 DB 566 GATATCATTCAGAGATCCCAAGAAATCCGATCGACATGACAGAGTTCACATTAGAG 625  
 QY 186 LeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIleLeuGlyAspTyrSer 205  
 DB 626 GGTGGCTCAAGCAATTTCTTTTCCCTCTCATGAGCCCTTCAGGCGCTCCATAG 685  
 QY 206 IleGlnValGlnValAsnAsp-----GlnThrTyrTyrGlnSerPheGlnValSerGlu 223  
 DB 686 GTGGTGCTACAGAAAGAAATCAGGTGAGAGACAGACAC--CCTTTCACCGTGGAGAGA 742  
 QY 224 TyrValIleProLysPheGluValThrLeuGlnThrProLeuTyrCysSerMetAsnSer 243  
 DB 743 TTGTGTTCTCCCAAGTTGATGATGACAGTAAACAGTCCCAAGATTAATCCATCTTGGA 802  
 QY 244 LysHisIleLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProValIleLysGlyAsp 263  
 DB 803 GAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 862  
 QY 264 ValThrLeuThrPheLeuProLeu----- 271  
 DB 863 GTGACTGTGACATTGTCAGAAAGTATGATGACGCTTCGACCTGCCACGCGTGAAGATTCA 922  
 QY 272 -----SerPheTyrGlyLys 276  
 DB 923 CAGGCTTTCTGTGAGAAATTCAGTGCAGCTAAACAGCCATGCTGCTTCATACAGCA 982  
 QY 277 LysLysAsnIleThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAsp 296  
 DB 983 GAAAAA-----ACCAAGCTCTTCACGCTGAGAGAG-----AAGGAG 1018  
 QY 297 GluGluMetLysAsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSer 316  
 DB 1019 TATGAATGAAA-----CTTCACACTGAGGCCACCATGCCA 1054  
 QY 317 SerProGlyProValGluIleLeuThr-----ThrValThrGluSerVal 331  
 DB 1055 GAAAGAGAAACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1114  
 QY 332 ThrGlyIleSerArgAsnValSerThrAsnValPhePheGlnHisAspTyrIleIle 351

DB 1115 ACCAAACCTCTCA---TTTGTGAAGTGACCTCACACTTTCAGACAGGGA-----ATT 1162  
 QY 352 GluPhePheAspTyrThrThrValLeuLysProSerLeuAsnPheThrAlaThrValLys 371  
 DB 1163 CCTTCTTTGGG-----CAG 1177  
 QY 372 ValThrArgAlaAspGlyAsnGlnLeuThrLeuGluGluArgArgAsnValValIle 391  
 DB 1178 GTGCCCTAGTATGATGAGAAAGCGCTCCCTATACA-----AATTAAGTCA 1225  
 QY 392 ThrValThrGlnAsnThrTyrGluTyrTyrSerGlySerAsnSerGlyAsnGlnLys 411  
 DB 1226 TTATCC--AGAGAAATGAAGCAAACTATTAAC-----TCAATGCTACACGAGATGAG 1276  
 QY 412 MetGluAlaValGln---LysIleAsnTyrThr-----ValProGlnSerGlyThrPhe 428  
 DB 1277 CATGCCCTTGTGATGCTTCTATCAACCAACCAATGTTATGAGTACCTCTTATACGTTT 1336  
 QY 429 LysIleGluPhePro-----IleLeuGluAspSer 438  
 DB 1337 AGGCTCATTTACAGAGATCGATGCCCTGTTACGCTTACAGTGGGTCCAGAAAGAAC 1396  
 QY 439 SerGluLeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSer 458  
 DB 1397 GAAAGGACATCAACACTGTTAT----- 1420  
 QY 459 LeuPheLysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLys 478  
 DB 1421 CTGTGTTCTCCCAAGCAAGAGCTTTGTCCACCTTCCAGCCCATGCTCATGAACTACCC 1480  
 QY 479 ValGlySer-----ProPheGluLeuValIleSerGlyAsnLys-----Arg 492  
 DB 1491 TGTGCCATCTACAGACAGTCCAGGACACATTAATTTGAAATGAGAGCAACCTGCTGGGG 1540  
 QY 493 LeuLysGluLeuSer-----TyrMetValIleSerArgIleGlnLeuValAlaValGly 510  
 DB 1541 CTGAAGAACTCTCTCTCTATATATGATGATGATGATGATGATGATGATGATGATGATGAT 1600  
 QY 511 -----LysGlnAsnSerThr-----MetPheSerLeuThr 520  
 DB 1601 ACTCATGACCTGCTGTGTAACAGAGACATGAAAGGCCATTTTTCATCTCAATCCCT 1660  
 QY 521 ProGluAsnSerTyrThrProLysAlaCysValIleValTyrTyrIleGluAspGly 540  
 DB 1661 GTGAAGTGAACATGCTCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1720  
 QY 541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560  
 DB 1721 GACGTGATTTGGGAGATTCGCAAAATATGATGATGATGATGATGATGATGATGATGATGAT 1780  
 QY 561 LeuTyrTyrSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580  
 DB 1781 TTGAGCTTCAAGCCCATCAAAAGCTCCCAAGCTCCACAGCCCACTGCGAGTACAGCG 1840  
 QY 581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600  
 DB 1841 GCT---CCTCAGTCCGTGTGCGCCCTCGTCTGTGAGCAAAAGCGTGTCTCATGAG 1897  
 QY 601 AlaSerAsnAspIleThrMetGluAsnValValIleGluLeu-----GluLeuTyrAsnThr 619  
 DB 1898 CCGATGTGAGCTCTCGCGCTCTGTTTCAACCTGCTTACCAAGAAAGAGACTCACT 1957  
 QY 620 GlyTyr-----TyrLeuGlyMet 625  
 DB 1958 GCGTTCCTCGGCGCTTGAATGACAGAGCATGATGATGATGATGATGATGATGATGATGATGAT 2017  
 QY 626 PheMetAsnSer-----PheAla 631  
 DB 2018 TATATTAATGGAATCACTATATCTCCAGTATCAAGTACAAATGAAAGATATGTACAGC 2077  
 QY 632 ValPheGlnIleCysGlyLeuTyrPheValLeuThrAspAlaAsnLeuThrLysAspTyrIle 651

Db 2078 TTCCTAGAGACATGGGCTTAAGGCAATTCACCAATCCAAAGATTCTGTAACCCCAAAATG 2137  
 Qy AspglyValTyrAspAsnAlaGlu-----Tyr 660  
 Db 2138 TGTCCACAGCTTCAACAGTATGAAATGACGTAGACCTCAAGCTTACGCTGATGTTTAT 2197  
 Qy AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680  
 Db 2198 GAGTCAGATGTAATGGGAAGAGCCATGACGCTGGTGCATGTTGAAGAG----- 2248  
 Qy GlycerSerProHis-----ValArgLysHisAspPheProGluTyrTrpIleTyrLeu 697  
 Db 2249 -----CCTCACACGAGACCGGTACGAAAGTACTTCCCTGAGACATGGATCTGGGAT 2299  
 Qy AspThrAsnMetGlyTyrArgIleTyrGlnGluPheGluValTrpValProAspSerIle 717  
 Db 2300 TTGGTGGCTGTAATCTCAGCAGGTGTGCTGAGTGAAGATGAACAGTCCCGAACCATC 2359  
 Qy ThrSerTrpValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThr 737  
 Db 2360 ACCGAGTGAAGACGAGGGGCTTCTGCTGTGTAAGATGCTGGACTGTGATCTCTCC 2419  
 Qy ThrProValGluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSer 757  
 Db 2420 ACT---GCCCTCTCCGAGCCTTCCAGCCCTTCTTGTGAGACTCACAAATCCTTACTCT 2476  
 Qy ValIleArgGlyGluGluPheAlaLeuGluTyrIleThrIlePheAsnTyrLeuLysAspAla 777  
 Db 2477 GTGATTCGTGAGAGGCTTCACTCACTCAAGGCCACGGCTCTAACTCTTCCCAATGC 2536  
 Qy ThrGluValLysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSer 797  
 Db 2537 ATCCGGGTCACTGTGACAGCTGGAAGCCTCCCGCTTCTTACGCTGCCAGTGAGAG 2596  
 Qy Glu-----IleAsnAlaThrGlyHisGlnGlnTrpLeuValProSer 812  
 Db 2597 GAACAAAGCCCTCACTGCATCTGTGCAACGCGCGCA----- 2635  
 Qy GluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGlyGluIleProIle 832  
 Db 2636 -----ACTGTCTCTGGGACGTAAACCCAAAGTCAATTAAGAAATGTAAATTTC 2683  
 Qy ThrValThrAlaLeuSer-----ProThrAla 841  
 Db 2684 ACTGTAGCGCAGAGGACCTAGATCTCAAGAGCTGTGGAGCTGAGTCCCTTCAATT 2743  
 Qy -----SerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIle 856  
 Db 2744 CCTGAACAGGAGAGAAAGACACAGATCATCAAGCCTGTGTGTGAACCTGAAGAGACTA 2803  
 Qy GluLysSerTyrSerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThr 876  
 Db 2804 GAGAAAGAAACAACTTCACTCCCTACTTGTTCATCAGTGGTGGAGTTTC----- 2857  
 Qy LeuLysThrLeuSerPheSerPheProProAsnThrValThrLysSerGluArgValGln 896  
 Db 2858 ---GAGAATATTATCCGTAACCTGCACCAATATGGTAGAAGAAATCGCCGACTTCT 2914  
 Qy IleThrAlaIleGlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArg 916  
 Db 2915 GTCTCAGTTTGGAGACATATTAGGCTCTGCATGCAAAACACAAAAATCTTCTCAG 2974  
 Qy MetProTyrGlyCysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeu 936  
 Db 2975 ATGGCCCTATGGCTGGAGACCAAGAAATATGGTCTCTTGTCTCTTAACATATATGACTG 3034  
 Qy AspTyrLeuThrLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPhe 956  
 Db 3035 GATTATCTTAATGAACACAGACCTTACTCAGAGGTCAAGTCCAAAGCCATTGGCTAT 3094  
 Qy MetArgGlnGlyTyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAla 976  
 Db 3095 CTCAACACTGGTATCCAGAGACAGTTGAACAAACACTATGATGCTCTTACAGACACC 3154

Qy PheGly---AsnTyrAspProSer---GlySerThrTrpLeuSerAlaPheValLeuArg 994  
 Db 3155 TTTGGGAGCGAATATGGACAGAAACCAAGGCGCAACCTGGCTTACAGCTTGTTCGAAG 3214  
 Qy CysPheLeuGluAlaAspProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyr 1014  
 Db 3215 ACTTTTGCCCACTGAGGCTGAGCTCATCTTCATGATGAAGACACATTTACCAAGCCCTC 3274  
 Qy ThrTrpLeuLysGlyIleGlnLysSerAsnGlyGluPheThrAspProGlyArgValIle 1034  
 Db 3275 ATATGGCTTCCCAAGAGCGAAGAGCAATGCTGTTCACGAGCTCTGGGTCACTGCTC 3334  
 Qy HisSerGluLeuGlnGlyIleAsnLysSerProValThrLeuThrAlaTyrIleValThr 1054  
 Db 3335 AACAAAGCCATTAAGGAGAGAGATGAAGATGAAGTGAACCTTCGCCGTATATCACCATC 3394  
 Qy SerLeuLeuGlyTyrArgLysTyrGlnProAsnIleAspValGlnGluSerIleHisPhe 1074  
 Db 3395 GCCCTCTGGAGATTCCTCTCACAGTCACTACCCCTGTGTCCGCAATGCCCTGTTTGC 3454  
 Qy LeuGluSer-----GluPheSerArgGlyIleSerAspAsnTyrThr 1088  
 Db 3455 CTGAGTCACTCTGGAAGACACACAAAGAGGGGACCATGGC---AGCCATGATATACC 3511  
 Qy LeuAlaLeuIleThrTyrAlaLeuSerSerValGlySer---ProLysAlaLysGluAla 1107  
 Db 3512 AAAGCACTGTGGCCATGCTTTGGCTTGGCCGAGGATACCAAGACAAAGAGAAAGTA 3571  
 Qy LeuAsnMetLeuThrTrpArgAlaGluGlnGlyGlyMetGlnPheTrp----- 1124  
 Db 3572 CTCAACTCATCTATGAGAGGAGTGAAGAAAGCAACTGTCTCCATTGGAGCCCTC 3631  
 Qy ValSerSerGluSerLysLeuSerAspSerTrpGlnProArg-----SerLeuAspIle 1142  
 Db 3632 CAGAAACCAAGGACCAAGTGGGCAATTTTACAAACCCAGGCTCTCTGCTGAGGTG 3691  
 Qy GluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGln-----ThrSer 1159  
 Db 3692 GAGATGACATCTCATGTGCTCTCGCT---TATCTCAGGCGCCAGGACGAGCCCACTCG 3748  
 Qy GluGlyIlePro-----IleMetArgTrpLeuSerArgGlnArgAsnSerLeu 1175  
 Db 3749 GAGGACCTGACCTCTCAACCAACATCGTGAAGTGAATCAAGAAAGCAGAAATGCCAG 3808  
 Qy GlyGlyPheAlaSerThrGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAla 1195  
 Db 3809 GCGGTTTCTCTCCACCCAGACAGACAGTGTGCTCTCCATGCTGTCCAATATGCA 3868  
 Qy AlaLeuMetAsnThrGluArgThrAsnIleGlnValThrValThrGlyProSerSerPro 1215  
 Db 3869 GCAGCCACATTTAC---AGGACTGGGAAGGCTGCACAGGTGACTATCCAGTTCACGG 3925  
 Qy SerProLeuAlaValValGln----- 1222  
 Db 3926 ACATTTTCCAGAAATTCAGAGTGACAACAACACCGCTGTACTGACAGACTGTCA 3985  
 Qy -----ProMetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCys 1238  
 Db 3986 TTGCCAGAGCTGCTGGGGAATACAGCATGAAGTGAACAGGAAGAGTGTCTACCTC 4045  
 Qy GlnLeuAsnValValTyrAsnValLysAlaSerGlySerSerArgArgArgSerIle 1258  
 Db 4046 CAGACATCTCTGAATATCAATAT-----CTC 4072  
 Qy GlnAsnGlnGluAlaPheAspLeuAspValAlaValLys-----GluAsnLysAsp 1275  
 Db 4073 CCAGAAAGAAAGATGTTCCCTTGTCTTAGGTGACAGACTGCTGCTCAAACTTGAT 4132  
 Qy AspLeuAsn---HisValAspLeuAsnValCysThrSerPheSerGlyProGly----- 1292  
 Db 4133 GAACCCAAAGCCCAACAGCTTCCAAATCTCCCTTAAGTGTCAATTACACAGGAGCCGC 4192

QY 1293 ---ArgSerGlyMetAlaLeuMetGluValAsnLeuLeuSerGlyPheMetValProSer 1311  
 Db 4193 TCTGCTCCACATCGCATCGCTGATGGAAGATGCTCTGCTTCAATCCCTGAAG 4252  
 QY 1312 GluAlaIleSerLeu-----SerGluThrValIleValGluIlyrAspHisGly 1328  
 Db 4253 CCACACGTAAGAAATGCTTGAAGATCTAACATGTGAGCCGACAGAAATCCAGCAGCAAC 4312  
 QY 1329 LysLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCysValAsnIleProAla 1348  
 Db 4313 CATGCTTGATTTACTTCACTGATAGGTGTCGAATCAGACACTGAGCTTGTCTTCACGGT 4372  
 QY 1349 ValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSerIleValAspTyr 1368  
 Db 4372 CTGCAAGATGTCCTCCATGAAGATCTGAACCAAGCAGCATAGTAAGTATGATTAATCTAC 4432  
 QY 1369 GluProArgArgGluAlaValArgSerTyrAsnSer 1380  
 Db 4433 GAGACGATGAGTTGCAATTCCTGATGATCAATGCT 4468  
 RESULT 15  
 US-09-756-247-3  
 ; Sequence 3, Application US/09756247  
 ; Publication No. US20030180722A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Godbole, Shubhada D  
 ; APPLICANT: Boyle, Bryan J  
 ; APPLICANT: Mize, Nancy K  
 ; APPLICANT: Deng, Cenhua  
 ; APPLICANT: Goodrich, Kyle  
 ; APPLICANT: Arterburn, Matthew C  
 ; APPLICANT: Zhou, Ping  
 ; APPLICANT: Tang, Y. Tom  
 ; APPLICANT: Liu, Chenghua  
 ; APPLICANT: Yeung, George  
 ; APPLICANT: Drmanac, Radoje T  
 ; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO ALPHA-2-MACROGLOBULIN-LIKE PC  
 ; FILE REFERENCE: HVS-31CIP  
 ; CURRENT APPLICATION NUMBER: US/09/756,247  
 ; CURRENT FILING DATE: 2001-01-08  
 ; PRIOR APPLICATION NUMBER: 09/649,167  
 ; PRIOR FILING DATE: 2000-08-23  
 ; PRIOR APPLICATION NUMBER: 09/540,217  
 ; PRIOR FILING DATE: 2000-03-31  
 ; PRIOR APPLICATION NUMBER: 09/684,711  
 ; PRIOR FILING DATE: 2000-10-06  
 ; PRIOR APPLICATION NUMBER: 09/560,875  
 ; PRIOR FILING DATE: 2000-04-27  
 ; PRIOR APPLICATION NUMBER: 09/496,914  
 ; PRIOR FILING DATE: 2000-02-03  
 ; NUMBER OF SEQ ID NOS: 41  
 ; SOFTWARE: Patentin version 3.0  
 ; SEQ ID NO 3  
 ; LENGTH: 4527  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)..(4527)  
 ; US-09-756-247-3  
 Alignment Scores:  
 Pred. No.: 5,56e-133 Length: 4527  
 Score: 1438.00 Matches: 428  
 Percent Similarity: 44.54% Conservative: 274  
 Best Local Similarity: 27.16% Mismatches: 557  
 Query Match: 19.57% Indels: 317  
 Db: 13 Gaps: 46  
 US-10-020-095-4 (1-1428) x US-09-756-247-3 (1-4527)  
 QY 20 LeuAlaValAlaProGly-----ProArgPheLeuValThrAlaProGly 34

Db 28 TTGGCCCTATACACCGCATTTGACAGAGAACTTCCAACTACTGCTGATGACATTACAGGCC 87  
 QY 35 IleIleArg---ProGlyGlyAsnValThrIleGlyValGluLeuLeuGlnHisCysPro 53  
 Db 88 CGGCTAAATTTCCCTC-----TCCGTCAAGAGTTTGTGTTGACCTGAGCCCT 135  
 QY 54 SerGluValThrValIleValAlaGluLeuLeuLysThrAlaSerAsnLeuThrValSerVal 73  
 Db 136 GGGTACACTGATGATTTAAATTACGGTTACTCTGGAGACCAAGAGCAACAGCAGAGTTG 195  
 QY 74 LeuGluAlaGluGlyValPheGluLysGlySerPheLysThrLeuThrLeuProSerLeu 93  
 Db 196 CTAGATATCTGTGGACTGAGAGAGGACCTTACTATATCTCTTCTTGTACACACT 255  
 QY 94 ProLeuAsnSerAlaAspGluIleTyrGluLeuLysValThrGlyArgThrGluAspGlu 113  
 Db 256 CCGTCTGTGGACAGAGAAAGTGGCCAAATCCGGGTGTGGGA--GTTGAAATATAC 312  
 QY 114 IleLeuPheSerAsnSerThrArgLeuSerPheGluThrLysArgIleSerValPheIle 133  
 Db 313 ATCAGCTTTGACGAGAAAGAAAGTTCTAATTCAGAGGACAGGGAACGACCTTTGTA 372  
 QY 134 GlnThrAspLysAlaLeuTyrLysProLysGlnGluValLysPheArgIleValThrLeu 153  
 Db 373 CAGACTGACAAACCTCTTACACCCAGGAGAGAGTATTTCCGATGTCACCAAG 432  
 QY 154 PheSerAspPheLysProTyrLysThrSerLeuAsnIleLeu---IleLysAspProLys 172  
 Db 433 GATAGCAACTTCCTGTCAGTAATGACAAAGTACTCCAGTGTGGAACTACAGATTCCAAT 492  
 QY 173 SerAsnLeuIleGlnGlnThrPheLeuSerGlnIleAsnAspLeuGluValIleSerLysThr 192  
 Db 493 AGCAACAGATTTGCACAGTGGCTGGAAGTGTACTGACAGCAGGACATTTGACCTGTCC 552  
 QY 193 PheGluLeuSerSerHisProIleLeuGlyAspTyrPheIleGlnValGluValAsnAsp 212  
 Db 553 TTCCAACTGACACCAAGAGCAATGCTGGACCTACACT-----GTGGCAGTGGCTGAG 606  
 QY 213 GlnThrTyrTyrGlnSerPheGlnValSerGluTyrValLeuProLysPheGluValThr 232  
 Db 607 GGCAGAACCTTTGTGATCTTTCAGTGTGAGAAATATGCTGCCCAAGTTTAAGGTGANA 666  
 QY 223 LeuGlnThrProLeuTyrCysSerMetAsnSerLysIleAsnGlyThrIleThrAla 252  
 Db 667 GTGGTGAACCCAGAGATTTATCAACGTGCAGGAATTTCTTAGTAAATTTGTTGT 726  
 QY 253 LysTyrThrTyrGlyLysProValLysGlyAspValThrLeuThrPheLeuProLeuSer 272  
 Db 727 AGGTACACCTATGGAAGCCCATGCTAGGGCAGATGATCTGTGTGT----- 777  
 QY 273 PheTrpGlyLysLysAsnIleThrLysThrPheLysIleAsnGlySerAlaAsnPhe 292  
 Db 778 -----CAGAAAGGCAAT-----ACTTAC 795  
 QY 293 SerPheAsnAspGluGluMetLysAsnValMetAspSerSerAsnGlyLeuSerGluThr 312  
 Db 796 TGGTATCGAGAGGTGAACGGGAACGCTTCTCTGCAAAATGACAGAACCTCTTGAGACAG 855  
 QY 313 LeuAsp-----LeuSerSerPro----- 318  
 Db 856 ACTGACAAACAGATGTTTCTACGACCTGTGACATGGCCACCTTTGACCTCATTTGGA 915  
 QY 319 -----GlyProValGluIleLeuThrThrValThrGluSerValThrGlyIle 334  
 Db 916 TATGCGTACAGCCATCAATCAATATTTGTGCTACTGTTGGAGAGGACAGGTGTG 975  
 QY 335 SerArgAsnValSerThrAsnValPhePheLysGlnHisAspTyrIleIleGluPhePhe 354  
 Db 976 GAGGCCAATGCCACTCAGAAATATCTACATTTTCTCCAAATGGATCAATGACCTTTGAA 1035  
 QY 355 AspTyrThrThrValLeuLysProSerLeuAsnPheThrAlaThrValIleValThrArg 374



Db 1036 GACACCAACAATTTTACATCAAAATTTCCCTTCAGTGGGAAGATAGATTAGGGCC 1095  
 Qy 375 AlaAspGlyAsnGlnLeuThrLeuGluGluArgAsnValValIle----- 391  
 Db 1096 CATGATGATCTCTCCTC-----AAGAACCATCTAGTGTTCGTGTGATT 1140  
 Qy 392 -----ThrValThrGlnArgAsn----- 397  
 Db 1141 TATGACCAAAATGGAACCTTCACACGACCTGCTACTGATTAACAATGGCTAGCTCCC 1200  
 Qy 398 -----TyrThrGlnTyrTrpSerGlySerAsnSerGlyAsnGlnLysMetGluAla 414  
 Db 1201 TTATACCTTGAGACATCCGGTTGGTAATGGACAAAC----- 1236  
 Qy 415 ValGlnLysIleAsnTyrThrValProGlnSerGlyThrPheLysIleGlu----- 431  
 Db 1237 -----GTTTCTCTGGAGGGAAGTTTCAAAATGGAAGAACTTAGTA 1275  
 Qy 432 -----PheProIleLeuGlnAspSerSerGlyLysLeuGlnLysAlaTyr 446  
 Db 1276 TATATCCGGAACAAGTCCAGCTTACTACCAAAATGCTACTGACCTGCACTCCGACCTTC 1335  
 Qy 447 PheLeuGlySerLysSerSerMetAlaValHisSerLeuPheLysSerProSerLysThr 466  
 Db 1336 TACAGCAACAACCCGACCTTCCTGGCATCCACCGGCTA----- 1374  
 Qy 467 TyrIleGlnLeuLysThrArgAspGluAsnIleLysValGlySerProPheGlnLeuVal 486  
 Db 1375 -----AACGGCCCTTGAAGTGGCAAGTGGCCAGCCCAAGGAAGTGTCTG 1413  
 Qy 487 Val-----SerGlyAsnLysArgLysLeuLysSer 497  
 Db 1414 GTGATTATTATCATCGACCCGCGCATGACAGCCCTGACCAAGATC--AGCTTCTCC 1470  
 Qy 498 TyrMetValValSerArgGlyGlnLeuValAlaValGlyLysGlnAsn----- 513  
 Db 1471 TACTATTATTAAGGAAGAGTTGGTGTATGAGAGGGCGAAGAACATCGAACTTAAG 1530  
 Qy 514 -----SerThrMetPheSerLeuThrProGluAsnSerTrpThrPro 527  
 Db 1531 AAGAAAGCATGAAAGCTCTCTCTCTCTCATCTGACCTTCACTTGGAGATGGCCCT 1590  
 Qy 528 LysAlaCysValIleValTyrTyrIleGluAspAspGlyGluIleIleSerAspValLeu 547  
 Db 1591 GATCCTTCCTCGTGTATGATGATGATTTTCCCAAGTGAAGGTGTGATGACAAAT 1650  
 Qy 548 LysIleProValGlnLeuValPheLysAsnLysIleLysLeuTyrTrpSerLysValLys 567  
 Db 1651 CAGTTCTCAGTGGAGATGTGCTTGTGACATCAGGTTTCCCTTGGCTTCTCCCTCCAG 1710  
 Qy 568 AlaGluProSerGluLysValSerLeuArgIleSerValThrGlnProAspSerIleVal 587  
 Db 1711 CAGCTTCCAGAGCAGAGAGTGAAGTGCAGCTGCAGCAGCT---CCCGATCTCTGTGT 1767  
 Qy 588 GlyIleValAlaValAspLysSerValAsnLeuMetAsnAlaSerAsnAspIleThrMet 607  
 Db 1768 GCGCTCGGGGGGTGATGATGATGTCTTACTGCTTAGGCCA----- 1809  
 Qy 608 GluAsnValValHisGluLeuGlnLeuTyrAsnThrGlyTyrTyrLeuGlyMetPhe--- 626  
 Db 1810 -----GACAGAGAGCTGAGCAACCGCTCTGTCTAT---GGATGTTTCCA 1851  
 Qy 626 ----- 626  
 Db 1852 TTCTGTATGTCTCACTACCTCATCAAGTGGCTGATGATCATGCTCCAGTCTTGGC 1911  
 Qy 626 ----- 626  
 Db 1912 CCATGGAGACTTTCCTCAGCCCTCATTTGACCAATGCCCAAGGGCAATGGAGCAGAGCT 1971  
 Qy 627 -----MetAsnSerPheAlaValPheGlnGlu 635  
 Db 1972 TCCATTATCTGGAGGCGCTCGTCTCTGAAAGCAGCAGCACTTTTCAGCTTTTCCGGGAC 2031

Qy 636 CyeGlyLeuTrpValLeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyr 655  
 Db 2032 GTGGGCTGAAAATACTGCTCAATGCAAAATACAGAGCA---GTAGTTGCTCAGTAC 2088  
 Qy 656 AspAsnAlaGluTyrAla-----GluArgPheMetGlu 666  
 Db 2089 AGATCTCCAGAAATACAGCACTGCTATGGTGGCGGTGCTATCATCAAGAGGCTTTGAGTCA 2148  
 Qy 667 GluAsnGluLysIleValAlaPheIleAspPheSerLeuGlySerSerProHisVal 686  
 Db 2149 TCAACTCCTTATCATCAAGCAGAG-----GATTCTCAGTTC 2184  
 Qy 687 ArgLysHisPheProGluThrTrpIleTrpLeuAspThrAsnMetGlyTyrArgIleTyr 706  
 Db 2185 CGCAGATCTTCCAGAGACCTGCTGCTGGATCTGTTCTTATGTGTAATCTGGGGAG 2244  
 Qy 707 GlnGluPheGluValThrValProAspSerIleThrSerTrpValAlaThrGlyPheVal 726  
 Db 2245 GAGCGGTCCACGTCAAGTTCCTGACCGCATCACCGAGTGAAGCGATGAGTCTGTC 2304  
 Qy 727 IleSerGluAspLeuGlyLeuThrThrThrProValGluLeuGlnAlaPheGln 746  
 Db 2305 ACTTCCAGTCAGAGAGCTTCGGGCTTTCACCCACT---GTTGACTTACTGCTTTCAG 2361  
 Qy 747 ProPhePheIlePheLeuAsnLeuProTyrSerValIleArgGlyGluGluPheAlaLeu 766  
 Db 2362 CGGTTCTTTGTGACCTGACTCTCCCTTACTGATGTCGTGGGGAATCCTTTCGCTT 2421  
 Qy 767 GluIleThrIlePheAsnTyrLeuLysAspAlaThrGluValLysValIleIleGluLys 786  
 Db 2422 ACTGCCACATCTTCAATTAACCTTAAGAGATTCATCAGGCTTTCAGTCCGCTAA 2481  
 Qy 787 SerAspLysPheAspIleLeuMetThrSerSerGluIleAsnAlaThrGlyHisGln 806  
 Db 2482 TCGCATGATGACGACTGAGATATGATGCGCATTC-----CAGACC 2523  
 Qy 807 ThrLeuLeuValProSerGluAspGlyAlaThrValLeuPheProIleArgProThrHis 826  
 Db 2524 TCCAGTGTCTGTGTGTATGACGCAAAACCAACCATCGAATCATCAGCTGTCAAA 2583  
 Qy 827 LeuGlyGluIleProIleThrValThrAlaLeu----- 837  
 Db 2584 TTGGGTCACTAATTAACCTTACTATGATACAAAGATTCGACAGCAATGAACATGTGG 2643  
 Qy 838 -----SerProThrAlaSerAspAlaValThrGlnMetIleLeu 850  
 Db 2644 GGCAGAGAGGGTTTGTCCCAAAAGGCCGGAAGTGAACGCTCATCAAGCAGATCTC 2703  
 Qy 851 ValLysAlaGluGly-----IleGluLysSerTyrSerGlnSerIleLeuLeuAspLeu 868  
 Db 2704 GTCAACCTGAGGAGAGCTCGTGTGAGAGACACACAGC-----TCATTG 2748  
 Qy 869 ThrAspAsnArgLeuGlnSerThrLeuLysThrLeuSerPheSerPheProAsnThr 888  
 Db 2749 CTGTGCCAAAGAAAGAGGCGCATGCTGATCTGTCTCCCTGAGCTCCACAGTGCATAT 2808  
 Qy 889 ValThrGlySerGluArgValGlnIleThrAlaIleGly----- 901  
 Db 2809 GTTCTGACTCGACCAAGAGCTTATGTACGTTCTGTGGTAAAGCATTTAGATTTCTTAC 2868  
 Qy 902 -----AspValLeuGlyProSer 907  
 Db 2869 TCAGAAAGAAAGCGAATGAGAGCTGCAAGAGTGTGAGAGACATTTAGGCACAGCC 2928  
 Qy 908 IleAsnGlyLeuAlaSerLeuIleArgMetProTyrGlyCysGlyGluGlnAsnMetIle 927  
 Db 2929 CTCAGAACTGTGATGTGTGTGAGATGCCAGATGCGTGTGGCAGCAGAACATGTGTC 2988  
 Qy 928 AsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThrLysLysGlnLeuThrAsp 947  
 Db 2989 TTGTTGCTCCCATCATATATGCTTTCAGTACTGTGAGAAAGCAGGCGCTGTGACGAG 3048

QY 948 AsnLeuLyGluValAlaLeuSerPheMetArgGlnGlyTyrGlnArgGluLeuTyr 967  
 Db 3049 GAGATCATGCTCTGAGGAGTTCCTGGAAATAGGGTACAGAGAGCTGATGATTC 3108  
 QY 968 GlnArgGluAspGlySerPheSerAlaPheGlyAsnTyrAspProSerGlySerThrTrp 987  
 Db 3109 AAACACAGAAATGGCTCATACAGTGGCTTTGGGACCGAGATGGAAATGAAACACATGG 3168  
 QY 988 LeuSerAlaPheValLeuArgCysPheLeuGluAlaAspProTyrIleAspIleAspGln 1007  
 Db 3169 CTGACAGCGCTTGTCTCAAAATGCTTGGCCAAAGCTCAGAAATTCATCTTCATGATGCC 3228  
 QY 1008 AsnValLeuHisArgThrTyrThrTrpLeuLyGlyHisGlnLySerAsnGlyLeuPhe 1027  
 Db 3229 AAGAACATCAGAGATGCTCTCAAGTGGAGGAGAAACAGGTCCCGCTGCTCTAT 3288  
 QY 1028 TrpAspProGlyArgValIleHisSerGluLeuGlnGlyAlaAsnLySerProValThr 1047  
 Db 3289 GCCAACGTGGAAATGCTCTTCACACAGCTATGAGGGTGTGATGATGATGAGCTCC 3348  
 QY 1048 LeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArgLySerGlnProAsnIleAsp 1067  
 Db 3349 TTGACGTGCTATGTCACAGCTGATTCGAGAGATGGAAAGATGATGATGATGATGATG 3408  
 QY 1068 ValGlnGluSerIleHisPheLeuGluSerGluPheSerArgGlyIleSerAspAsnTyr 1087  
 Db 3409 GTGAGTCAGGCTCTGAGTGTCTCAAGATTCG---GCCACCTCCACAGCAACCTCTAC 3465  
 QY 1088 ThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGlySerProLyAlaLySerGlu--- 1106  
 Db 3466 AACAAGGCTCTGAGTGTCTTACATTTCTCTGCTGGGAAATGACATCAGAAACAT 3525  
 QY 1107 AlaLeuAsnMetLeuThrTrpArgAlaGlnGlnGlyGlyMetGlnPheTrpValSer 1126  
 Db 3526 CTCCTTAACAGATTAGATCAACAGCTATCATCTCAGAGAAATCATTACTGAGCCAG 3585  
 QY 1127 SerGluSerLyLeuSerAspSer---Trp---GlnProArgSerLeuAspIleGlu 1143  
 Db 3586 AAACCTACTCCATCATCGAACCCAGCCCTGCTGAGCTGCGGCTGATGATGAGAA 3645  
 QY 1144 ValAlaAlaTyrAlaLeuLeuSerHis---PheLeuGlnPheGlnThr 1158  
 Db 3646 CTCACAGCATATGATGTTGGCCAGCTTACCAAGCCAGCTGATCAAAAAGAGATA 3705  
 QY 1159 SerGluGlyIleProIleMetArgTrpLeuSerArgGlnArgAsnSerLeuGlyPhe 1178  
 Db 3706 GCGAAGCCACATAGCATATGCTGCTGGCCCAACACACATCATATGGGGCTTC 3765  
 QY 1179 AlaSerThrGlnAspThrThrValAlaLeuLyAlaLeuSerGluPheAla---Ala 1196  
 Db 3766 TCTTCTACTCAGGATATCTAGTGTCTTCAAGCTCTTGCCAAATATGCCACTACGCC 3825  
 QY 1197 LeuMetAsnThrGluArgThrAsnIleGlnValThrValThr--- 1210  
 Db 3826 TACATGCCATCTGAGGAGATCAACCTGTTGAAATTCACAGAAATTTCCAGCGCACA 3885  
 QY 1211 ---GlyProSerSerProSerProLeuAlaValValGln---Pro 1223  
 Db 3886 TTCACATATCAGTCACTTAAACATTTGTAATTCACAGAGATACCTGCCCAATGTCCCT 3945  
 QY 1224 MetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValVal 1243  
 Db 3946 GGAATGTACACGTGAGGCTTCAGGCCAGGGCTGTGTATGTGCAGACGGTGTGAGA 4005  
 QY 1244 TyrAsnValIleAlaSerLySerSerArgArgArgSerIleGlnAsnGlnGluAla 1263  
 Db 4006 TACAATATCTCCCTCC---ACAAATATGAAAGACC 4038  
 QY 1264 PheAspLeuAspValAlaValLyS---GluAsnLySAspAspLeuAsnHis 1279  
 Db 4039 TTTAGCTTAGTGGAATAGAAAAGCTAGATGTGAGCAACCACTTACCTCGATCC 4098  
 QY 1280 ValAspLeuAsnValCysThrSerPheSerGlyPro---GlyArgSerGlyMetAlaLeu 1298

Db 4099 TTGACTCTCATATTACACACGATTATGTGGGAGCCGTAGCTCTTCCAAATATGGCTATT 4158  
 QY 1299 MetGluValAsnLeuLeuSerGlyPhe---MetValProSerGluAlaIleSerLeu 1316  
 Db 4159 GTGGAAGTGAAGATGCTATCTGGGTTCACTAGTCCCATGAGAGGACCAATCAGTACTTCTC 4218  
 QY 1317 SerGlu---ThrValIleLyValGluTyrAspHisGlyLyLeuAsnLeuTyrLeuAsp 1335  
 Db 4219 CAGCAACCCCTGGTGAAAGGTTGAATTTGGAACCTGACACACTTAACATTACTTGGAT 4278  
 QY 1336 SerValAsnGluThrGlnPheCysValAsnIleProAlaValArgAsnPheLyValSer 1355  
 Db 4279 GAGCTCATTAAGACACATCAGACTTACCTTACCATCAAGCAAGTGTGCTGCTCACCC 4338  
 QY 1356 AsnThrGlnAspAlaSerValSerIleValAspTyrTyrGluProArgArgGlnAlaVal 1375  
 Db 4339 AACTTGAAACACGACCAATCAGGCTATAGTACTACTTACCA----- 4383  
 QY 1376 ArgSerTyrAsnSerGluValIleLyLeuSer-SerCysAspLeuCys 1390  
 Db 4384 -----GTTCTTTTAATTATCTCAGTACACAAATTGTGT 4417

Search completed: January 17, 2004, 08:49:35  
 Job time : 1238 secs

**THIS PAGE BLANK (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2004, 18:10:30 ; Search time 28 Seconds

(without alignments)  
4904.603 Million cell updates/sec

Title: US-10-020-095-4

Perfect score: 7348  
Sequence: 1 MGGPPLTAAHDLVCVTAL.....HSSVIFFCFKLLYFMEWL 1428

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1854.5	25.2	1508	2 T27828	hypothetical prote
2	1850	25.2	1539	2 T27829	hypothetical prote
3	1638	22.3	1507	2 T18544	alpha-2-macroglobu
4	1451	20.3	1476	2 JCS143	alpha-1 proteinase
5	1455	19.8	1477	2 A29952	alpha-1 proteinase
6	1440	19.6	1474	1 MAHU	alpha-2-macroglobu
7	1424	19.4	1503	2 T43166	alpha-2-macroglobu
8	1417	19.3	1487	2 S15904	alpha-1 proteinase
9	1412.5	19.2	1472	2 A26122	alpha-2-macroglobu
10	1402	19.1	1451	2 B41185	pregnancy zone pro
11	1396.5	19.0	1482	2 S13495	alpha-2-macroglobu
12	1395.5	19.0	1476	2 A41185	alpha-2-macroglobu
13	1395	19.0	1473	2 A20872	ovostatin precursor
14	1384	18.8	1500	2 A42210	alpha-1-macroglobu
15	1374.5	18.7	1464	2 JCS144	murinoglobulin pre
16	1373.5	18.7	1495	2 S27001	alpha-2-macroglobu
17	1373.5	18.0	1699	2 T14074	complement compone
18	1307.5	14.7	1683	2 T30885	complement compone
19	1079.5	14.0	1673	2 T50806	complement compone
20	1026.5	13.5	1652	2 T50711	complement C3 prec
21	987	13.4	1738	1 A24558	complement C4 prec
22	986.5	13.4	1735	2 A29176	sex-limited protei
23	982.5	13.4	1735	2 S54784	complement C4A pre
24	982.5	13.4	1744	1 C4HU	complement C3 prec
25	973.5	13.2	1651	1 C3NJ	complement C3 prec
26	966	13.1	1663	1 C3RT	complement C3 prec
27	963	13.1	1620	2 A29953	alpha-1 proteinase
28	960	13.1	1620	2 T51339	complement compone
29	956	13.0	1663	1 C3MS	complement C3 prec

30	934	12.7	1663	1 C3HU	complement C3 prec
31	921	12.5	1676	1 C5HU	complement C5 prec
32	905	12.3	1680	1 C5HU	complement C5 prec
33	904	12.3	1666	1 C3GP	complement C3 prec
34	894	12.2	1642	2 T51018	cobra venom factor
35	857.5	11.7	1620	2 S21045	complement protein
36	736.5	10.0	609	2 A41081	alpha-1-inhibitor
37	694	9.4	1230	2 T30517	complement C4 - Q2
38	688.5	9.4	1617	2 T28153	complement C4 - ch
39	467.5	6.4	1536	2 E72310	hypothetical prote
40	432.5	5.9	726	2 A27602	complement C3 - ra
41	290.5	4.0	1892	2 C97804	hypothetical prote
42	287	3.9	1653	2 B91052	hypothetical prote
43	287	3.9	1653	2 F85896	hypothetical prote
44	283	3.9	1653	2 G65028	hypothetical prote
45	269.5	3.7	1644	2 AC0823	probable lipoprote

## ALIGNMENTS

## RESULT 1

T27828 hypothetical protein ZK337.1a - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 17-Mar-2000

C/Accession: T27828

R/White, S.

submitted to the EMBL Data Library, November 1996

A/Reference number: Z20426

A/Accession: T27828

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-1508 <MIL>

A/Cross-references: EMBL:Z82090; PIDN:CA80506.1; GSPDB:GN00019; CESP:ZK337.1a

A/Experimental source: clone ZK337

C/Genetics:

A/Map position: 1

A/Intons: 36/2; 116/3; 201/1; 267/2; 338/3; 413/3; 527/3; 556/3; 691/3; 714/1; 760/2; 8

C/Superfamily: alpha-2-macroglobulin

Query Match 25.2%; Score 1854.5; DB 2; Length 1508;

Best Local Similarity 31.8%; Pred. No. 6.3e-97;

Matches 476; Conservative 292; Mismatches 562; Indels 169; Gaps 44;

QY	15	VCTAALAVAGPRPLVTAPGILIRPGANVTIGVLLHGPSQYTVKAEILKTASNLTVSVL	74
DB	28	VSTTAAPVKA-TYMLVAPAVVRPPQPSVCNMLKQATDEDMIVRIEVRTERNETIAR	86
QY	75	EAEGVFEKGSFKTLPLPLNSAD-EIYELRVGRQD-EILFNSNTRLSFEKRTISV	131
DB	87	VISNI-KPGIAQTVSLSEMPASLTIPROSRYKLYINGETLMAELIFENENELKYDOKALSV	145
QY	132	PLTQDKALYKPKQEVKFRIVTLFSDPKPKYKTSNLILIDPKSNLLOQMLSQSDIGVISK	191
DB	146	PLTQDKALYKPKQEVKFRIVTLFSDPKPKYKTSNLILIDPKSNLLOQMLSQSDIGVISK	205
QY	192	FTQLSHPLIGDWSIQV---VNDQYYQSPQSEVYLPRKPVTLQTPLYCSMNSKHLN	247
DB	206	ELQLABETLLGWMFLFEVERISNGVQDS---STVDYVLPKREVNIMKSSFTIND-DLS	261
QY	248	GTTTAKTYTGKRVKQDVTLTF-LPLSPF-----GKKK-----ITKPKIN--GS	289
DB	262	VFDVAKTYTGKRVKQDVTLTF-LPLSPF-----GKKK-----ITKPKIN--GS	321
QY	290	ANFSFNDEEMKVMDSNGLSEYLDLSSPGVEILTVTESVTGSRNVSTVFFKOHY	349
DB	322	AAVFSNDELK-----HKLHEMGGSIRIYASTEDTELEARNATHOISTPREEV	373
QY	350	IIFFDYTVTLKPSINFTATYKVRADGN--QTLLEBRNNVITVTOANTYVSGNS	407
DB	374	KIDVERKQDTPFKRGLTYNVVVALKQMDTPVYKATLPKR---VQSTFYNY--YNHDS	427

```

Qy 408 GNOKMEAVOKINTYVPOSGTFKIEPILEDSELOKAYFLGSKSMAVHSLFK----- 461
Db 428 SLOGEKRTKIYEVDAHGTSVLTLOPINCTSARIEAH-YDIGKDNFTATPISLYVEA 486
Qy 462 --SPSKTYIOUKTRDEN-IKVGSPEFLVVSNGKRLKELSYMVVSROLV-----AVGKON 513
Db 487 AVSPKTSFLQLLADNEGAVDVGKLSFSLKATOPSLITYOVMSRSNIIVSQOMTVNSEH 546
Qy 514 STMSFLTPENSWTPKACVIYV-YIEDDEGIISDVLIKIPVOLVFNKIKLYMSKYKAPSE 572
Db 547 ATI-SFPATANMAPKSKRLIVYALIESQEVLDALDFVEGIFQONVALSIDKQAVEFGQ 605
Qy 573 KVSRLISVTPDSIVGIVADVKSVMNANASNDITMENVHLELYNTGYLYGMF----- 626
Db 606 NVKFKVT-SDKNSFVGLLVVDQSVLLKTKGNDITREKVEQOLENVDNSNNVGGGFGGRPW 664
Qy 627 -----MNSFAVQEGCLMVLTDANLTD-----YIDGYIDNA 658
Db 665 EAIIDRRKRSIWRPMWIGIGSDAQSIIFSNAGLVLTDLALYREPOREFMSVMMMDGARGMA 724
Qy 659 EYARFPMENEGHIVDHFSLGSS---PHYRKHPPEPMWLTDTNMGRIYQEEVATVPD 715
Db 725 EAA--FAAPMG-----GSSPPPTVRKFPFPHTWMSDLN-STSGVEWEIEAPD 771
Qy 716 SITSMVATGVISEDGLGLTTTPEVLEQAFQPFIFLNLPSYVIRGEEFALEITIFNYLK 775
Db 772 TITSMVASTFAINENGLGVAPTTSKLVRFPFIQNLPLVAVRGEKFAVLVVFNYME 831
Qy 776 DATEVKYII--EKSDKFDILMTSSEI--NATHQOITLVSEDCAT--VLPIRPTH 827
Db 832 KEQOVTVTLKYDKOSGDLLKKDGTIVARDEVGOONVAIVSAGGTSKAYVFIIVPSI 891
Qy 828 GEIPIVTVALSPTASDAVTOMILVKAEGIEKSYQSILDLTDNRLOSTLTLTSPSPPN 887
Db 892 GEIPIHSAISOGSDAVENMLRDPQGYKVDKRNIPFIIDLNNSSDPS-KNLELIMPND 950
Qy 888 TVTSSERVQITAGDVLGPSINGLASLIRMEYGGCEOMNINFAPNIYILDYITKKQLTD 947
Db 951 VVDSQKARLDVIGMGMGPVNLNNAHKLVQMPYGGEGOMLNLVNLVVKTLRATNNES 1010
Qy 948 NLKEKALSFMKQGYORELLYQREDGSAFGNYDPSGSTMVSAFVLCFLEADPYIDIDQ 1007
Db 1011 QLETKAIKFIQOGLQRELTYKRADNSFSAFGSDSKAGSTMVLAFAVVSFHHAKQYAFADP 1070
Qy 1008 NVLHRTYTWLKGHOKSNGEPMWDPGRVHSELOGG-NKSPVLTAYITSLIGYKKYQNI 1066
Db 1071 NVISNAVAFLLNSQQMESGAFERGBVHHKMDQGAQDGGVALTAFAVLISIL-----ENGW 1125
Qy 1067 DVQESIHFLBSEFSGISDNTYLLALITYALSSVGSFKAKELANMLTW--AEDEGGMQFW 1124
Db 1126 ENGRAVYTLLEKHLDVEVSNATYMAVVALOLAKSKQAKAFENLKKHKIYKSGDVKFA 1185
Qy 1125 VSSP--SKLSDS---WQPSLDIEVAAYALLSHFLQFQISEGIPIMKWSLRQNSIGGF 1178
Db 1186 SAQKKEKVEKESRAVYMFQARPVDIETTSYAVLSYLAQWQTSSESIIIRMLVSGNEBGGF 1245
Qy 1179 ASTODTTVALKALSEBALMTERTNIQYTV-TGSPSSP-----PLAVVQPM----- 1224
Db 1246 TSTDYTWALQALSSYAAVITYSDHTSQVITLKNKHTHSFDINRMAIVLOSQYSLSLND 1305
Qy 1225 AVNISANGFGPAICQLNVNVVNVKASGSSRRRSIONCEAPDLDAVAKENKODLHNVLDN 1284
Db 1306 AVSINANGTGVYKQLSYSY-----YDSDLANDAPFCSEITEIRAG-NRLQDL 1355
Qy 1285 CTSGSGRSGMALMEVNLISGFVNPSEALSSETVK-----KYVYDHGKINLYLDSVNE 1339
Db 1356 CCNTRYRCKSNMALAEIDALSgyRFDQVHTLTSIEDLQVEMEKODTKMNVYFENPLGG 1415
Qy 1340 TQFCVNIIPAYANFKYSNTQDASVSIIVDYEPBRQAVNSYNSEVLSGCDLSDVQGRP 1398
Db 1416 RPYCLTSYSDVTVYVADQKRPANFRLVYDYDEEQLKMTYAKQOTRSLQEKCG--EDCW 1472

```

```

RESULT 2
T27829
hypochemical protein ZK337.1b - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000
C/Accession: T27829
R/White, S.
submitted to the EMBL Data Library, November 1996
A/Reference number: Z20426
A/Accession: T27829
A/Status: preliminary; translated from GB/EMBL/DDBI
A/Molecule type: DNA
A/Residues: 1-1519 <MIL>
A/Cross-references: EMBL:Z282090; PIDN:CAB05007.1; GSPDB:GN00019; CESP:ZK337.1b
A/Experimental source: clone ZK337
C/Genetics:
A/Gene: CESP:ZK337.1b
A/Map position: 1
A/Insertions: 36/2; 116/3; 201/1; 267/2; 338/3; 413/3; 527/3; 556/3; 691/3; 714/1; 771/2; 84
C/Superfamily: alpha-2-macroglobulin

```

```

Query Match 25.2% Score 1850; DB 2; Length 1519;
Best Local Similarity 31.6%; Pred. No. 1,1e-96;
Matches 477; Conservative 290; Mismatches 563; Indels 180; Gaps 44;

```

```

Qy 15 VCTAALAVAGPREFVTAPGIIIRPGGNVTIGVELLEHCHESQVTKAELKTASNLTVSVL 74
Db 28 VSTTAAPVKA-ITMLVAPAVAPDPQFSCNMNLQADDEDMIVALEVTEFNEETIAAR 86
Qy 75 EABGVKESGFKTLPLSLPLNSAD--EYELRVTRTOD-EILFNSRSLSEYTRISV 131
Db 87 VISNL-KPGIAQVTSISEMPOQSLTPROSKYKLIIRGETINAEILFENNELKQDQALSV 145
Qy 132 FIQTDLAKYKQEKVFRIVTLFSDPKPYTSINILIKQKSLIQOMLSQSDLGVISK 121
Db 146 FIQTDRAIKRPASLVVRALVKSILKPYGNATIKIPPSRNLISQITGLVLDRGVSG 205
Qy 192 TPOLSHPILGDSIOVO-----VNDQTYVQFOVSEVLPKFEVTLQTPLYCSMNSKHLN 247
Db 206 ELQLAEBETLGDWFIETVNSGVQDS---SFTVDVTVLPKFEVNIKTSSFITIND-DLS 261
Qy 248 GTTATKYTGPKYKGVNTLTF-LPLSFW-----GKKKN---ITKTKIN--GS 289
Db 262 VFVDAKTYGKGYAKAKVSLLEPMHRMAMVPTIIDENGVKKEELMVERTVKLNROGE 321
Qy 290 ANFSFNDDEKKNVMSDNGSEYLDLSSPGPVILTTVYESVYGISRNVSNTVPFQHDY 349
Db 322 AAVVFSNDELK-----HKLHEWGGGSIIRIVASVTEIDIEIRNATHQISTREEV 373
Qy 350 IIEFFDYTVLKPSLNFATVKTTRADGN--QLTEERRNNVITVQRYTEYMGSGNS 407
Db 374 KLVVEKQDIFKGLYNNVVVALKQMDDPVKATILKR-----VOVTFNYP--YVHDS 427
Qy 408 GNOKMEAVOKINTYVPOSGTFKIEPILEDSELOKAYFLGSKSMAVHSLFK----- 461
Db 428 SLOGEKRTKIYEVDAHGTSVLTLOPINCTSARIEAH-YDIGKDNFTATPISLYVEA 486
Qy 462 --SPSKTYIOUKTRDEN-IKVGSPEFLVVSNGKRLKELSYMVVSROLV-----AVGKON 513
Db 487 AVSPKTSFLQLLADNEGAVDVGKLSFSLKATOPSLITYOVMSRSNIIVSQOMTVNSEH 546
Qy 514 STMSFLTPENSWTPKACVIYV-YIEDDEGIISDVLIKIPVOLVFNKIKLYMSKYKAPSE 572
Db 547 ATI-SFPATANMAPKSKRLIVYALIESQEVLDALDFVEGIFQONVALSIDKQAVEFGQ 605
Qy 573 KVSRLISVTPDSIVGIVADVKSVMNANASNDITMENVHLELYNTGYLYGMF----- 626
Db 606 NVKFKVT-SDKNSFVGLLVVDQSVLLKTKGNDITREKVEQOLENVDNSNNVGGGFGGRPW 664
Qy 627 -----MNSFAVQEGCLMVLTDANLTD-----YIDGYIDNA-----NLT 647
Db 665 EAIIDRRKRSIWRPMWIGIGSDAQSIIFSNAGLVLTDLALYREPOREFMSERRLNTPGLT 724

```



```

Db 936 GNOGYSTVITADADAATROLVLEBEGPKEKEDMTSTFACPDQNGKFT-ATSULLLEBEDVED 994
Qy 892 SERVQITAIQDVLGSPSINGLASLIRMPYCGEONMINFAPNIYILDYITKKQJOTDNLKE 951
Db 995 SARGVSTGTQDLMGPALKNLDHLVRLPTGCGEONMVKFVPNIPLVDLYITATGTSITDSIKE 1054
Qy 952 KALSPMQGYQRELLYQREDEDSFSAFGNYDPDSSGTWLSAFLRCLFEADPIPIDQONVH 1011
Db 1055 KALNNMRGVARQONRYRHPDGSYSAFGNRDQXGMLFLTAFYRSPFAQERFILNKKKN 1114
Qy 1012 RTYTMKLGKHSQNEFMDPGRVHSELOCG----NKSPVLTATVIVSLT--GYRKQPN 1065
Db 1115 ETEMWILNRQDSNGCFKRTIGKLFNSALKGKISSNDETPAPLTAVYLSLEAGYKA---- 1170
Qy 1066 IDVOESHFLHESEFSPRGIS-----DNYLTALITYALSSVGSPPKAEKALNMLTWPRAQ 1117
Db 1171 -----ETVIDQGISCLEALSNPSTYSIALFAATSLAGHPAKNDYLAKEBRAIT 1220
Qy 1118 ECGMQFWVSSSKLSDSQPRSLDIEVAANY--LSHFLOFQTSSEGIPIKRWLSRQNS 1174
Db 1221 ECGRTFPMKSPSS--GRYYWGNISIGVETAGYAVLTLLHGGSANLAKTPIIRMLAKOONY 1278
Qy 1175 LGFSASTODTIVALKALSEFALNTEBRTNIQVYVGTGSSSPSLAVUQ----- 1222
Db 1279 RGGYSTVDITVIALIQAMSKFATITIKOBLDEVGESSGFEKKMILTQNSILMOTFRLO 1338
Qy 1223 --PAVAINISANVGFGAICQLNVVNVNVKASGSSRRRSRISQNOEAPDLVAVKEN--KODLNL 1278
Db 1339 TVBSPVDEALAGSGCGLVQTSIRYNVNP-----PPRKGHLBYTVYKRGJLYRCIN 1389
Qy 1279 HVDLNVCTSPSGPGR-SCMALMEVNLISGFVWPSBAIS-----LSFTVKVYVDHGKLT 1332
Db 1390 -AH1ATCYKYGDKGGSNMVAILEMKMVSQWIPDEESIKNIIVDBELNLRREYVDQGNLT 1448
Qy 1333 YLDSVNEHQFCVNIPAVKNFKYSNTQDASVSVIVYIERRQAVASVNSYEVKLSQCD 1388
Db 1449 YFSELTQDNLCFNFWMLEQDIEVOETKPAITILYLYVELEQDEVVNSYSID--ENCCE 1501

```

```

RESULT 4
JC5143
alpha-macroglobulin precursor - guinea pig
C/Species: Cavia porcellus (guinea pig)
C/Date: 31-Jan-1997 #sequence_revision 27-Feb-1997 #text_change 20-Jun-2000
C/Accession: JC5143
R/Iwasaki, H.; Suzuki, Y.; Sinohara, H.
J. Biochem. 120, 1167-1175, 1996
A/Title: Cloning and sequencing of cDNAs encoding plasma alpha-macroglobulin and murinog
A/Reference number: JC5143; MUID:97164019; PMID:9010766
A/Contents: liver
A/Accession: JC5143
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-1476 <IWA>
A/Cross-references: DBJ:D84338; NID:G1805591; PIDN:BA12316.1; PID:G1304084
C/Comment: This protein belongs to alpha 2-macroglobulin family.
C/Superfamily: alpha-2-macroglobulin
C/1-23/Domain: signal sequence #status predicted <SIG>

```

[illegible]

Db	120	VLVSKXSLVFNQVDRKTYRGGQVVARVUSLDENFPR-----LNELPPLFIQDPKGNRV	175
Qy	177	QOMLSQSDJGVISKTFOJSSHPILGDMSTQV-QVNDQTYQSFQVSEYVLPRFEVTLQOT	235
Db	176	MOQONLKLERGLTQJSPPLSSEPLILGSGSVVHKESGGRMHSPVLEEPLPKFEVQVSM	235
Qy	236	PLYSNMNSKHLNGITAKTYTGKPVKGDVLTFL-----PLSTWKKKKNTKTF--KING	288
Db	236	PKKJITILEQEFVTSVCGRYTGGKVPNGPNTWMSICRNYNNSACLSSE--SRAFCCKYNQ	292
Qy	289	SAN-----FSPNDEEMK-----NVMDSNLSLEVLDDJSSGPPEILITVT	328
Db	293	QJNSQGFQIQVKTNDQJLRKREYEMKLREAKIREGTC-----VOLTGCFSEITATIT	348
Qy	329	ESVYTGIRSNVSTNVFPKQHDYIIIEFPDYTLVRLPSLNFATVAVTRADGNQJLTLEERRN	388
Db	349	K-----LSFVAVDSYVRBGVFFQGVRL--VDGKNVPMPHM--	383
Qy	389	VVITVQNTREYVSGNSGNQKMEAV-KINTVPSGCFPKIEFLLEBSS----ELQJ	443
Db	384	ITITASEKNY-----HSNATTDEGLVQFSINTTNMIGTSLNTQVKHKDSTNCYQJWLL	438
Qy	444	KAYFLGSSMAVASHLSPKSPKTYIOJKT-----RDNINKVGSPEFLVSGNKRJKE	495
Db	439	EANSGASHTANAVSL-----SRSGVHLPEPQJLPCHQJOTQFKA---HJILKG-QJLKE	489
Qy	496	L--SYMVVSRGQJLAVG-----KONST-----WESLTPENSWTPKACVIVYIYEDDEIT	543
Db	490	LVFVYVIMAKGVIQSGYVYLVSQJNTKGHFVSIVSVESDLPVAVARVILYAILPSGEIL	549
Qy	544	SDVLKIPQVLVFKKKIKLJWSKYKAESSEKYSLRISTQDSDJGVIVAVKSNVL-----	598
Db	550	ADSAKYNVENCLDNKKNVLSFSEGSJLSPASXTHLRVTAS--POSICLAPAVDQSVLLRKPEA	608
Qy	599	-MNASN-----DIT-----MENVYHELENTGYGYGMF-----MNSFA	631
Db	609	VJASSSVYALLPVKDLTGFPCLLQOQEBNQECSVLNT-YIDJLSYSPENINEXDMYG	667
Qy	632	VFOECGLMWLNDANLTQDYIDGVYDNEVAREFMEENEGHIVDHFSLGSSPH-----	685
Db	668	FLXOMGLKVFNTKTIQKQJCAHVQKEVPLMAYSSESS-----SPRSGPRVAVG	720
Qy	686	-----VAKHPPEYIW--LDTNMGRHIOYEFVYVYPDJSTTSVAVGCVJSED	730
Db	721	IAATYSEBPKEATVATYSPEIWMJLKLKJTDSSG--VAEVEVYVDTTEWKAGAFCLSD	777
Qy	731	LGJLGTTPVLOAFOPPEFLFJLMPSYVINGEPEALTEITFNYLKDAETVKVJIEKSKF	790
Db	778	TGLGJLSPT-ASLRAFQPPVELTJMPYSIVIGEATTLAVALNTLPDJCIRLSVHLEASPKE	836
Qy	791	DILMTSEINATGQOJTLVPSBEDGAVLFPPIRFTLGEJPIVTA-----	836
Db	837	LAERKAE-----QESGCVCGNRRQIVSWVVPKJSGNNPFTYSANALSSSELGNEKT	890
Qy	837	LSPT--ASDAVTOMILVYAKGIEKYSQCSILLDTNRLQSTLKTJSSFPENVTJGSE	894
Db	891	VVPTYGKKDITIKKPLVPEEGIEKEEFMTSLIRVSDTJVS---EKJLLELPSWVJQDSAR	947
Qy	895	VQJTRAGVUGPSJINGLASLIRMYGGGGEOMNINFAINIVILDYLTCKKJLONTLKEKAL	954
Db	948	ATVSIILGJIGJASAMQNTQNLQOMVYGGGEOMVWFAENIVYLDJLNTQJLTPJISKAI	1007
Qy	955	SFMQSGVORELLYREDGSSFAFG-AN-DESGSTWLSAFVJRCFLPADPYJIDQNTLHR	1012
Db	1008	SYLSTGYQJQJNTHGRBOSYTBEBENRGQGNWMLTAPVLKJTSQARKKIFIDEAHITQ	1067
Qy	1013	TYVJLKHQKNSGEFMDGPAVHISELQGNKSPVTLTAVYVTSILGYRKYQPNIDQEST	1072
Db	1068	ALSWLSQKQKNGCFWSSGSLNNNAIKGVEDETISAVYITIALLEMSLDPJTHPVANAL	1127
Qy	1073	HFLBS-----EFSRGJSDNTYLLALITYALLSVGS-PKAEALNMLTWRAEQEGMGQFWV	1125
Db	1128	FCLESAMWKAKEGTHG--SHVYTKJLALVAFALANQQRKEILKSLDEGVKENSJLHWA	1186





Db 1133 SSWENIEQGGNGSFVYTKALMAVAFALAGNQEKNELKSLDKAEIKEDNSIHMERPOKP 1192  
 QY 1125 VSSSEKLSDSWQPSRLDIEVAAYALLSHFLQFOTSEGP-----IMRMLSRORNS 1174  
 Db 1193 TKSGCYLYTP-QASSAEKEMSAIYVLAARL-----TAQPAPESEDIALSMGTTKMLTKOONS 1247  
 QY 1175 LGGFASITQDTTVALKALSER-AALMNTERTNIQVTVTPSSPSPLAVVQ----- 1222  
 Db 1248 YGGRSSITQDTVALDALSKYGATFESKSQKTPSTVVOSSGFSQKQFQVQDKSNRLLLQOVS 1307  
 QY 1223 ----PMAVNISANQGFRCIQLNVTNYVYKASGSSRRRRRSTIONQAPFLDPAVK---ENK 1274  
 Db 1308 LPYIPGVYTVSVSEGCYVAGTTIRYVPLE-----KQAPAFALKYQTVPLTCNPP 1358  
 QY 1275 DDLGHVDLVNCTSPSGP-GRSGMALMEVNLISGFVNSPESALISLETYKVV-----E 1324  
 Db 1359 KGQNSFQIISTEISTWGSRPASNNVIADKMLSGF-----TLPKPYTKKERLCHVSRTE 1412  
 QY 1325 YDHGKMLYLDVSVNEITPCVNIIPAVRNFKKVSNTODASVSIIVYVYEPFRQAVRSYNS 1380  
 Db 1413 VTTNNVLLYLDQVNTQNTLSFSEFIILQODIPVKNLQPAIVKYVDYETDEVAFAEYSS 1468

## RESULT 6

MAHU  
 alpha-2-macroglobulin precursor [validated] - human  
 N:Alternate names: alpha-2M  
 C:Species: Homo sapiens (man)  
 C:Date: 05-Apr-1993 #sequence-revision 30-Jun-1987 #text-change 08-Dec-2000  
 C:Accession: A94033; J19375; S09107; JN0262; A92486; S6634; A01256  
 R:Kan, C.C.; Solomon, E.; Balt, K.T.; Chalm, A.C.; Hiron, L.R.; Fey, G.  
 Proc. Natl. Acad. Sci. U.S.A. 82, 2282-2286, 1985  
 A:Title: Nucleotide sequence of cDNA encoding human alpha-2-macroglobulin and assignment  
 A:Reference number: A94033; MUID:85190481; PMID:2581245  
 A:Accession: A94033  
 A:Molecule type: mRNA  
 A:Residues: 1-1474 <MAN>  
 A:Cross-references: GB:M1133; NID:G177869; PIDN:AAAS1551.1; PID:G177870  
 A:Note: hydrolysis of the thiolester bond during amino acid sequencing of the mature pro  
 lymphomism  
 R:Beil, G.T.; Rall, L.B.; Sanchez-Pescador, R.; Merryweather, J.P.; Scott, J.; Bddy, R.L  
 Somat. Cell Mol. Genet. 11, 285-289, 1985  
 A:Title: Human alpha-2-macroglobulin gene is located on chromosome 12.  
 A:Reference number: 139375; MUID:85219061; PMID:2408344  
 A:Accession: 139375  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 832-999; 'I', 1001-1147; 'D', 1149-1194; 'D', 1196-1474 <BEL>  
 A:Cross-references: GB:M36501; NID:G177871; PIDN:AAAS1552.1; PID:G177872  
 R:Marynen, P.; Devriendt, K.; van den Berghe, H.; Cassiman, J.J.  
 FEBS Lett. 262, 349-352, 1990  
 A:Title: A genetic polymorphism in a functional domain of human pregnancy zone protein:  
 n.  
 A:Reference number: S09106; MUID:90242963; PMID:1692292  
 A:Accession: S09107  
 A:Status: not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 672-747 <MAR>  
 R:Matthijs, G.; Devriendt, K.; Cassiman, J.J.; Van Den Berghe, H.; Marynen, P.  
 Biochem. Biophys. Res. Commun. 184, 596-603, 1992  
 A:Title: Structure of the human alpha-2-macroglobulin gene and its promoter.  
 A:Reference number: JN0262; MUID:92246939; PMID:1374237  
 A:Accession: JN0262  
 A:Status: not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 1-29 <MAT>  
 R:Soltrup-Jensen, L.; Stepanik, T.M.; Kristensen, T.; Wierzbicki, D.M.; Jones, C.M.; Lot  
 J. Biol. Chem. 259, 8318-8327, 1984  
 A:Title: Primary structure of human alpha-2-macroglobulin. V. The complete structure.  
 A:Reference number: A92486; MUID:84239807; PMID:6203908  
 A:Accession: A92486  
 A:Molecule type: protein  
 A:Residues: 24-562; 'E', 564-974; 'Z', 976-999; 'I', 1001-1474 <SOT>

R:Soltrup-Jensen, L.; Stepanik, T.M.; Kristensen, T.; Wierzbicki, D.M.; Jones, C.M.; Lot  
 J. Biol. Chem. 260, 6500, 1985  
 A:Reference number: A92529  
 A:Contents: annotation; extratum  
 R:Virca, G.D.; Salveen, G.S.; Travis, J.  
 Hoppe-Seyler's Z. Physiol. Chem. 364, 1297-1302, 1983  
 A:Title: Human neutrophil elastase and cathepsin G cleavage sites in the bait region of  
 A:Reference number: A91713; MUID:84030513; PMID:6195065  
 A:Contents: annotation; inhibitory site  
 R:Soltrup-Jensen, L.; Lonbald, P.B.; Stepanik, T.M.; Petersen, T.E.; Magnusson, S.; Jor  
 FEBS Lett. 127, 167-173, 1981  
 A:Title: Primary structure of the 'bait' region for proteinases in alpha-2-macroglobulin  
 A:Reference number: A91290; MUID:81212827; PMID:6165619  
 A:Contents: annotation; inhibitory site  
 R:Hall, P.K.; Nelles, L.P.; Travis, J.; Roberts, R.C.  
 Biochem. Biophys. Res. Commun. 100, 8-16, 1981  
 A:Title: Proteolytic cleavage sites on alpha-2-macroglobulin resulting in proteinase binc  
 A:Reference number: A90099; MUID:81255805; PMID:6167263  
 A:Contents: annotation; inhibitory site  
 R:Mortensen, S.B.; Soltrup-Jensen, L.; Hansen, H.F.; Petersen, T.E.; Magnusson, S.  
 FEBS Lett. 135, 295-300, 1981  
 A:Title: Primary and secondary cleavage sites in the bait region of alpha-2-macroglobulin  
 A:Reference number: A91299; MUID:82056610; PMID:6172288  
 A:Contents: annotation; inhibitory site  
 R:Dolmer, K.; Jenner, L.B.; Jacobsen, L.; Andersen, G.R.; Koch, T.J.; Thirup, S.; Soltrup  
 FEBS Lett. 372, 93-95, 1995  
 A:Title: Crystallisation and preliminary X-ray analysis of the receptor-binding domain of  
 A:Reference number: S6634; MUID:9602553; PMID:7556651  
 A:Accession: S6634  
 A:Molecule type: protein  
 A:Residues: 1337-1343 <DOL>  
 C:Comment: This inhibitor is able to inhibit all four classes (EC 3.4.21-3.4.24) of endo  
 s tertiary conformation, entrapping the proteinase. This results in hydrolysis of its thi  
 e and appears to unmask a receptor-specific binding site.  
 C:Comment: The entrapped enzyme remains active against low molecular weight substrates (e  
 r complex results in its rapid clearance from the circulation by receptor-mediated endoc  
 C:Comment: The wide specificity of this inhibitor is attributed to the primary sequence c  
 ities of a variety of plasma proteinases, form primary (residues 704-809) and secondary  
 C:Generics:  
 A:Gene: GDB:A2M  
 A:Cross-references: GDB:119639; OMIM:103950  
 A:Map position: 12p13.3-12p12.3  
 C:Complex: homotetramer; dimer of disulfide linked dimers  
 C:Superfamily: alpha-2-macroglobulin  
 C:Keywords: glycoprotein; plasma; proteinase inhibitor; thiolester bond  
 F:1-23/Domain: signal sequence #status predicted <SIG>  
 F:24-1474/Product: alpha-2-macroglobulin #status experimental <MA2>  
 F:693-694,704-709,719-723,730-735/Region: inhibitory #status predicted  
 F:698-742/Region: bait region  
 F:48-86,251-299,269-287,278-431,595-771,642-689,821-849,847-883,921-1321,1079-1127,1352-  
 F:45,70-247,396-410,869,991,1424/Binding site: carbohydrate (Asn) (covalent) #status exp  
 F:410,563/Disulfide bonds: interchain #status predicted  
 F:972-975/Cross-link: thiolester (Cys-Gln) #status experimental

Query Match 19.6%; Score 1440; DB 1; Length 1474;  
 Best local similarity 29.1%; Pred. No. 2, 2e-73;  
 Matches 451; Conservative 266; Mismatches 551; Indels 284; Gaps 59;

QY 12 LHCCTAALAVADPPFLVTAPGRIIRPGGVNTIGVELHCSPQVTKAEELKTASNLTV 71  
 Db 15 LVLTLPTDASVSGKPPQVMVLVPSLHT-ETTEKGCVLISLNTVNTVSALESVRGNRL 73  
 QY 72 -SVLEAG-VFEKSGKRTILTPSLNSADE---IYELVYGTQDEILFNSNTRLSFER 126  
 Db 74 FTDLAEADVILHCVAF-----AVPKSSNNEVMTLVQVKGFTQE---FKKRTYVWVKN 124  
 QY 127 KRISVLEIOTDKALKYKKEQVFRIVTLPSDFPKYKTSNLTL-ISKDKSNLIQOMLSQOSD 185  
 Db 125 EDSLVVQIDKSIYKKGQYTKFRVSMDENFHNLNLIPLVYIQDPKGRIRIAQMSQGLE 184  
 QY 186 LGVISTPQLSSHPILIGDSIQOVND--QTYQSQVSESYLVPKREVTLQTPLYSGMNS 243  
 Db 185 GGIKQSPFLSSPFGSGYKVVQKKSGRTEH-PTVEEFVLPKFEVQVTVPKIITILE 243

[illegible]

```

QY 1176 GGFASDTDTVALKALSEFALNMTREPTNIQVTTGSSSPBLAVQ----- 1222
Db 1246 GGFSSDTDTVALHALSKYGAFTF-RTGKAQVTTSSGTFSSKPFVDNNRLLQOVS 1304
QY 1223 ----PMWNISANGFGFALCQLVNVNVKASGSSRRRBSRIONCAFDLDAVK--ENKD 1275
Db 1305 LPELPEXSMKVTBEGCVYLQTSIKNI-----LPEKEFPALGVQLPQICD 1353
QY 1276 DLN-HVDLNVCTSFSGF--RSGMALMEVNLISGFVWPSEALS--SETVKVEYDHG 1328
Db 1354 EPKATHSFQISLSTVSYGSRASNMALVDVGVSGFIPLKRTVXMLERSHNVSTREVS 1413
QY 1329 KLANLYDSVNETQECVNIIPAVNRKVSNTODASISYDYIEPRQAVRSYNS 1380
Db 1414 HVLITLKVSNQTLSEFTVYLQDVPRVHDLKPAIKVVDYETDEFAIAEYNA 1465

RESULT 7
T43166
alpha-2-macroglobulin - Japanese lamprey
C|Species: Lampetra japonica (Japanese lamprey)
C|Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 20-Jun-2000
C|Accession: T43166
R|Nonaka, M.
submitted to the EMBL Data Library, November 1992
A|Reference number: Z22323
A|Accession: T43166
A|Status: preliminary; translated from GB/EMBL/DBJ
A|Molecule type: mRNA
A|Residues: 1-1503 <N>
A|Cross-references: EMBL:D13567; PIDN:BAA02762.1
A|Experimental source: liver; clone Cl.3 and Cl.10
C|Superfamily: alpha-2-macroglobulin

Query Match 19.4%; Score 1424; DB 2; Length 1503;
Matches 442; Conservative 256; Mismatches 600; Indels 266; Gaps 48;

QY 10 AHLTCVCTAALAAVAGPR---PLVTPAGIIRPGANVTIGVELHPCSOVTVAEELKT 65
Db 7 AFLLLCALAAVADQSSGHYIVPVPSELHLSERLCVSL-CVTGEVTPPATLHYK 65
QY 66 ASNLTVSLAEAGVFEKGSFKTLTPS-----LPLNSADEIYELVGTGTOEI 114
Db 66 DSRDTSHVAQNHVTAVGEVDAGVLLQPSGGLDHCFSFTVP-DVQGTIVANLVRAAGBL 124
QY 115 LFSNSTLSFPTKRIISPIQTDKALYKPKQEVKRYITLSDPR-KYTSINLIKPKPS 173
Db 125 NFYTHAVVAKVADVVFQTDKVKYKPGOSVKFRVVLIDENFATVLTALYIEBDFOR 184
QY 174 NLIQOMLSQOSDLVICKTFLQSLSHPLIGMSIOV---QVNDQTYGOSFOYSEVVLPRFE 230
Db 185 NRIQOMNRASGRAGIVQELDMSEPLPGITVNVVWBOGSDSVASHFTVEEYVLLPFE 244
QY 231 VTLQTPLYCSNNSKHLNGITTAKTYYKPKVGDVTLFLPLSFWKKKKNTIKTFKINGSA 290
Db 245 VSIQTPSELVNLDSVTLKVCGRRTYKPKPHGAVNAVC-----IQGP 288
QY 291 NFSNDEBKKVMD-----SSNGLSEY-LDLSFG-----PELITTVTESVTGIS 335
Db 289 RFWRREERICIPVCEHFFPMKYKQKGCACAWOYNAPKPSNASCHTTHVLKVVALLBEGTG 348
QY 336 RNVTSTNVFFKHODYIIEFPDYTLAKSLNFTATVKTADAGNQLTEBRNNNVITVTO 395
Db 349 MKQDAEKNFEDITRISFVDMPSYRGLRPLGVKVKERPDGSAVP---HKLVSILYKQ 404
QY 396 RNYTEYWSGNSGNQKMEAVQKINVTYVPOSGTFKIEPIILDESDOLUKAYFLIGS---KS 452
Db 405 GNAAPRPSQNHGT-----ADGTFFETI-DIGDFNRSDTIFLEADNDFPNSPAHBSVTVQO 458
QY 453 SMAVHSLFKPSKSKYITOLKTRDENIKYGS--PEFLVSGNKLKE----- 495
Db 459 GYSLTISAFFPSDSEFLDIDRVANLLEGGSSVPLLELLV---LKENRSSAGDHAGVAVH 514

```

Qy	496	---LSTWVSRGOL-----	---VAVGKONSMTFSITPENSMTPRACVIVYIEDDGEII	543
Db	515	PIVINTLVMSRGNIIHTTESFTMDNYVLGKRAHSHHFOUKYHSAFPAIRKLAMVFNKN-ETV	573	
Qy	544	SDVLKIPVQVLFKNKIKLYMSKYKAEPSKVSIRISVTPDSDIVGIVAVDKSVNLMNASN	603	
Db	574	ADTLALPHAKCPFNKSVQFEGHEOELPGAMSTLIRAS-PSGLCALRVYDKSVLLKPEA	632	
Qy	604	DITMNVVHELELYNTGY-----	---LGMFNMSF-----AVQBEGL	638
Db	633	QLSASISIFORLOVODLTYDYDGIERWDCSPHYRRKRKSYGWMDDFONSVRLEPKGMGL	692	
Qy	639	WLTJANTLTKOYID-----	---GYDDNAEYAEPRMEENEGHIYDHDPSLGS--	683
Db	693	VVLDTTTV-KASVDCHQOLMHRHRIRIAYSMETVASFGLRPSDKG----	---NISPVGGGAAF	746
Qy	684	---PHVKHPPEETWIMLDTNMGYRIYQOFEVTVDSDITSWATGVISEDGLGLTTP	739	
Db	747	KTYETVAEYEPFEITWIDLYPSESGLEBVANAKVDSITEMQASAF-CSPAPGFLSVEVS	805	
Qy	740	VELQAFOPPFIFLNLPSYVIRGEPALETTINYLKDATEVYKIIIEKSDKPDILMTSBI	799	
Db	806	-SLRFTPEFVEPVLPSYSVRGETPEPLAISVANYLHSCIKIEVTLIDSEHFAY-----	857	
Qy	800	NATGHOQ---TLVSESDGATVLPRIPIRLHEIPIITYALS-----	---PT	840
Db	858	---AAGPGRGAASVCVCDRKG-LPSFLIEPLALGTAVNSPAAVAHSDPCQNEVVVPEBG	915	
Qy	841	ASDATONMILVACIEGIESYOSILIDLTDRNLOSTLKTLSFSPPNVTYSGERVQITAI	900	
Db	916	AVDTVRSYVIEBEGIPKELAYSSL-LCPKSPASERFVFNULPLRANVIVGSARAYATIA	973	
Qy	901	GDVLGSEINGLASLIRMPYGCGEQNMNIPAPNIYILDYLTKKKOLTDNLKEKALSFMROG	960	
Db	974	GDIMGSAIONDKLTLTLPFGCGEQNMVKEPAPNIYIOEYLONGQOITDAVRDALKALNFLRVG	1033	
Qy	961	YORELLYQREBDSFSAFGNDYPSGTSWLSAFPLRCFLEADPIYIDQNVLAHTTYMLKCH	1022	
Db	1034	YOROLTYRKRDHSYSAFGSDDDGQTMWLTALFALKSFVRASKHAIVSEDHITGPFMWLVEH	1093	
Qy	1021	OK-SNGEPMDDPRVHSELOGNKSQPVTLTAVITVSLGCRYKOYINID---VOESIH---	1072	
Db	1094	QNASITGCTISVGRLEFNNAKKGVSDDVSLITAYITALL-----ESNISGPVERALGCLR	1144	
Qy	1074	--FLESEFSRGISDNYTLALITYALS-SVGSPEKAEKALMMLTWRAEOBGMOFNWVSBSK	1130	
Db	1149	PLVLES-----SNHLLALASAYASLSADGATAGALRALQSRVATKGGLHMQEANK	1200	
Qy	1131	LSDSQWOPR-----SLDIEVAAYALLS-----HFLQFQISEGIPIMRWLSR	1170	
Db	1203	DDKDEDEEGENFRFSRYGTTSSAIVETTAYALLSRLAVPQGLAASATSSNIGVQWMLSK	1262	
Qy	1171	QRNSLIGSASTODTVALKALSEFAL-----	---MATERITNI-QUTYVGP	1210
Db	1263	QRNAAVGFSSTQDITVGLQALSAFALIDDGCGGSGGKAVLVDLSAHTLLREVSISTP	1322	
Qy	1213	SS-----PSPLAVQPMANISANGGFAICQLNVYVYNKASGSSRRRSRISQNEAF	1264	
Db	1323	NALIIHQVQLPRPPVAPAMVSSCTVATAGOGCALPQVSLKTIKEPPSSKPK-----F	1377	
Qy	1265	DLDVAVKE---NKODLNVHDLNVCTSFSGP-GRSGMALMEVNLISGFVWPVSEASISLSEN-	1310	
Db	1374	SLSVETNPANGSKAKARKGFQINVERYSYHGERGESNMALVEYVMISGYSAVKSLSKELQGF	1433	
Qy	1320	---YAKVEYDHGKMLYLDVSVNETQPCVNIIPAVNRPKYSNTODASVSYDYIYEPKRAVR	1376	
Db	1434	YDFYKRVAVDSDSRVVIYLDKVDKKETAIKLSVTODIAVDNLOPATVRVYDYATATEDAATS	1493	
Qy	1377	SYNS	1380	
Db	1494	MYSA	1497	

RESULT 8  
 S15904  
 alpha-1 proteinase inhibitor III, variant 1 precursor - rat  
 C|Species: Rattus norvegicus (Norway rat)  
 C|Date: 30-Jun-1992 #sequence revision 30-Jun-1992 #ext.change 16-Jul-1999  
 C|Accession: S15904; S52235; S29737; B34278; B29952; C29952; D29952; B27199  
 R|Regier, R.; Sicking, S.; Schweizer, M.  
 FEBS Lett. 282, 368-372, 1991  
 A|Title: Differential regulation of the two mRNA species of the rodent negative acute phase  
 A|Reference number: S15904; MUID:91243832; PMID:1709877  
 A|Accession: S15904  
 A|Molecule type: mRNA  
 A|Residues: 1-1487 <REG>  
 A|Cross-references: EMBL:X52984; NID:g55561; PIDN:CAA37176.1; PID:g55562  
 R|Sicking, S.; Schweizer, M.  
 submitted to the EMBL Data Library, May 1992  
 A|Reference number: S52235  
 A|Accession: S52235  
 A|Status: preliminary  
 A|Molecule type: DNA  
 A|Residues: 1-161 <SIG>  
 A|Cross-references: EMBL:X66454; NID:g609245; PIDN:CAA47070.1; PID:g609246  
 R|Thomsen, N.K.; Soltrop-Jensen, L.  
 Arch. Biochem. Biophys. 300, 327-336, 1993  
 A|Title: alpha-macroglobulin domain structure studied by specific limited proteolysis.  
 A|Reference number: S29737; MUID:93143334; PMID:7678727  
 A|Accession: S29737  
 A|Molecule type: protein  
 A|Residues: 25-32,576-280,'XG','314-320,'X','322','SL','717','VAV','722-726','Q',728,1205-1218  
 R|Englund, J.U.; Salvesen, G.; Thøgersen, I.B.; Plazzo, S.V.  
 J. Biol. Chem. 264, 11428-11435, 1989  
 A|Title: Proteinase binding and inhibition by the monomeric alpha-macroglobulin rat alpha  
 A|Reference number: A34278; MUID:89291900; PMID:2472336  
 A|Accession: B34278  
 A|Molecule type: protein  
 A|Residues: 684-715,'N','717-733 <ENG>  
 R|Brcicak, T.A.; Northemann, W.; Hudson, G.O.; Shleis, B.R.; Gehring, M.R.; Fey, G.H.  
 J. Biol. Chem. 263, 3599-4012, 1988  
 A|Title: Sequence and acute phase regulation of rat alpha-1-inhibitor III messenger RNA.  
 A|Reference number: A92709; MUID:88153707; PMID:2831216  
 A|Accession: B29952  
 A|Molecule type: mRNA  
 A|Residues: 1-3,'D','5-86,'V','88-136,'W','139-169,361-364,'W','366,'V','368-397,'H','399-727,'  
 A|Cross-references: GB:J03552; GB:M22359  
 A|Note: the authors translated the codon ATG for residue 174 as Ile  
 A|Accession: C29952  
 A|Molecule type: mRNA  
 A|Residues: 80-262,269-361 <BR2>  
 A|Cross-references: GB:M22360; NID:9202583; PIDN:AAA40633.1; PID:g554402  
 A|Accession: D29952  
 A|Molecule type: mRNA  
 A|Residues: 315-409,1064-1079,'TAAGAPDHC',1090,'TMP' <BR3>  
 A|Cross-references: GB:M22361; GB:M22362  
 R|Schweizer, M.; Takabayashi, K.; Geisler, T.; Laux, T.; Biermann, G.; Buhler, J.M.; Gauth  
 Eur. J. Biochem. 164, 375-381, 1987  
 A|Title: Identification and sequencing of cDNA clones for the rodent negative acute-phase  
 A|Reference number: A27199; MUID:87190405; PMID:2436907  
 A|Accession: B27199  
 A|Molecule type: mRNA  
 A|Residues: 1416-1483,'N','1485-1487 <SCH>  
 A|Cross-references: GB:M28297  
 A|Experimental source: liver, clone p13CDNA3  
 C|Genes:  
 A|Introns: 29/2; 90/3; 144/1  
 C|Superfamily: alpha-2-macroglobulin  
 C|Keywords: proteinase inhibitor; thiolester bond  
 F|1-24/Domains: signal sequence #status predicted <SIG>  
 F|25-1487/Product: alpha-1 inhibitor III variant 1 #status predicted <MAT>  
 F|648-723/Region: bait region #status predicted  
 F|985-988/Cross-link: thiolester (Cys-Gln) #status predicted

Best Local Similarity 29.0%; Pred. No. 4,5e-72; Matches 454; Conservative 244; Mismatches 576; Indels 290; Gaps 55;

13 LCVCSTALAVAPGPRFLVTAAGIIRPGN---VTGVLELHPCPSQVTKAEKLTAAEN 68  
 9 LCLFSLALFLFPASLL-----NNSKTMVLVPSQLYTETPEKCLHLIYHLETVT 59  
 69 LTVSVLEAG-----VEKGSFKTL--TLPSLPLNSADELYELRVGTODELLFSN 118  
 60 VTASLISQGTGKRLPDELVLVDKLFHCLSFITLPLPSSEESLDINIGAHK---PFE 116  
 119 STRLSETRISVFIOTDKALKYKQGVKFRIVTLFSDKPKYKTSLNIL-----IKPKS 173  
 117 RRVLVKNKESVVFQVTDKPVVPGQSVKFRVVMKLNHP-----LNEFLPLAYIEDPKM 172  
 174 NLIQGLSQQSDLGVSITKTFOLSHPILDMSI---QVAVDQVYQGSFOVSEVYLPRF 229  
 173 NRIWQODIKTEGKQLSFSLSAEPIQGVKIVILKQGVKEE--HSFTWMEFLPRF 229  
 230 EYTLQTPLYCSMNSKHLNGITITAKTYGKPVKGDVTLTLPL--PLSFMGKKINITYTKFI 286  
 230 GVDVKNPNAISVDELIINTVACALITYGKPVGHVAKISLCHGNP--SFSETSACKE--- 285  
 287 NSANFSFNDEKKNVMSNGLS-----EYLDSSPGPVEILTTTESVTG 333  
 286 -----EDSELDN-----NGCSTQEVNITEFOLKENYIKMHQ--AFHVNATVTEEGTG 330  
 334 ISRNVTNVPFQKHDIIEFFDVTYVLKPSLNPATAVKTRADGNQNLTEERNVTV 393  
 331 SEBSGGRLEVEETRNKPLFLKADSHFRHGIPFVKIRLVDIKGDIPNEQ---VFIRA 386  
 394 TORNTYEWGSGNSGQKMEAVOKINVTYVPGSTFKIEFPILEDSESLDKAVFLSGSKS 453  
 387 QGAGYTNATTTOHGLAK-----FSDTSSISGYS--LMIKITYH--KEES 427  
 454 MAVHS-----LEKSPSKTYIQLKTR-----DENIKVGSPPFEL--VVSQN 490  
 428 SCIHSSCTARRHAEBHTAVAVYSLSKYIYLDTEAGVLPQNIQHTVQAHFLIKGVLGV 487  
 491 KALKELSYWVVRGQLVANG-----QNSTMSLIPENSWTKACVIVYIYD 538  
 488 LPOIVFHYLYMAQGSILQTNHNTHOVEPGVSOVQGFALIEPFEFMVAVAKLITTLIP 547  
 539 DEEITSDVAKIPQVLFKRIKILYMSKVAEPESEKSLRISTVQPSIYIYAVDKSVNL 598  
 548 DGEVLADSTVFOVEKCLRNKHLSEFSSQSLPASQTHMKTAS--PQSLGGLNADVDQSVLL 606  
 599 MNASNDITMENVHEL-----EL--YNYGYVLYGMF 626  
 607 LKPEALFS--PSLIYDLPGMQDSNFISSVHPFEDEVDCIMYQPRDTEELTVGPV--GRE 663  
 627 NMSFAVFOEGCLMVLTDANL---TKY-IDGYDNAEYERMEENBEHGIYDIHDSLGS 682  
 664 KQYRVRVROMGLTAFNLKIKHPTCYEKMMVVLSPAVSESLSPRGSGF--EMMPLGVNK 722  
 683 SP-----HVKHPEPTIMVLDTMAGRIYQGFVTVYVDSITSVANG 724  
 723 SPLPKPPKDPKPPKQVLETIRNHPPEWIMDLVTNSSGTVTEVMTVPTITTEKKAQ 782  
 725 FVISEDLAGLTTTPEVLQAFPPFLFNLPSYVINGEFALEITITFNYIKDATEVKVIL 784  
 783 LCLSNDTGGLSIV--ATLQAFQFFVELIMPYSVINGEAFMLKATVMYTLPLSLPAVQL 841  
 785 EKSDKDIIMTSE-----INATG-HQOQLVPSSEGAIVLPI-----RPHHGEIPI 832  
 842 EASPDPTAVPVGDODSYCLGANGRHTSSMLVTPKSLGNVNSVSVEAQOOSPELCSQVA 901  
 833 TTTALPTASDAYTOMILVKAEGIEKYSOSILDLTDNRLOSTLTKLSFSPPTVYTS 892  
 902 TVPBTG--KRDIVYKVLIVPEPKIKHPTSSILCSDELFS---ETLSLLPPTVYKDS 956  
 893 ERVQITAGDVGLGSPINGLASLIRMPYGCGEQNMIFAPNIYILDYLTKKQQLTDNLKEX 952

Db 957 ARAHPSVNGDILSALKNTQNLQMPYGCGEQNMVLPAPNIYLVKYNETOQLTEKISK 1016  
 QY 953 ALSFMRQCYORELLAYOREDSFSAFNGYD--PSGSTWLSAFVLCFLEADYDIDQVNL 1010  
 Db 1017 ALGYIRAGYQREELNKHDSYSAFGDHNGGCGNTWLTALVLSFAQARAFIPIDESHI 1076  
 QY 1011 HRTYTMKGHQKNGEFPDGRVVIHSELQGNKSPVTLTAVIYVSLGKRYKOPNID--V 1068  
 Db 1077 TDAFTWLSKQKDCSCFRSSGLFNNMAKGVDEITLSAVITMALL--ESSLPDTBIV 1134  
 QY 1069 QESHFLES-----EFSRGISDNTTLLITLALSVS--PAKELANMLTRAEQEGMQ 1122  
 Db 1135 SKALGCEASWETIEQGNRSFVYTKTLMAYAFALAGQEKRENTLKSIDREAIRDMSI 1194  
 QY 1123 FM-----VSESEKLSDSMQRSLDIEVAVALSHFLQFQTSRGIP-----IMR 1166  
 Db 1195 HMERQKTKSEGYLYTP-QASSAEVESAIVVLAFL---TAQAPSPEDLALSMGTIK 1249  
 QY 1167 WLSQRNSLAGFPASTQDTTVALKALSEF--AALMNTERTNIQVTVGPSSPPLAVQ--- 1222  
 Db 1250 WLTQQNSHGFGFSSTQDTVVALDLASKYGAATFSKQKPLVTIQQSSGFSQKQVQNSN 1309  
 QY 1223 -----PMAVISANGFPAICQLNVNYYNKAAGSSRRRRSIQNEADLDVAV 1270  
 Db 1310 RLILQOVSLPDI PGNYTVSVSGEGCVVAQTLIRYMPLE-----KQOPAPALKVQT 1360  
 QY 1271 K-----ENKDDLNHDVNLVCTSPSGP--GRSGMALREVNLSGFVWPSAISLSEYTKVE- 1324  
 Db 1361 VPLTNKNGQNSFOISLEISYTSGRPASNVIVADVKMLSGF-----IPKPTVKKLER 1414  
 QY 1325 YDH-----GKNTLYDSVNETQPCVNIIPAVNFKVNSNTODASVSIYDVEEPQAVR 1376  
 Db 1415 LEHVSREVTNNVNLVLYDQYTNQTLSSFSFIQDIPKKNLQPAIVKYYDYETDEVAFA 1474  
 QY 1377 STNS 1380  
 Db 1475 EYSS 1478

## RESULT 9

A:6122  
 N:Alpha-2-macroglobulin precursor [validated] - rat  
 N:Alternate names: alpha-2M  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 16-Aug-1988 #sequence revision 16-Aug-1988 #text change 31-Dec-2000  
 C:Accession: A26122; A05278; S03843; A2614; I52403; I54013  
 R:Geierling, M.R.; Shleis, B.R.; Northemann, W.; de Bruijn, M.H.L.; Kan, C.C.; Chain, A.C.  
 J. Biol. Chem. 262, 446-454, 1987  
 A:Title: Sequence of rat liver alpha-2-macroglobulin and acute phase control of its mess  
 A:Reference number: A26122; MUID:87083488; PMID:2432068  
 A:Accession: A26122  
 A:Molecule type: mRNA  
 A:Residues: 1-1472 <GB>  
 A:Cross-references: GB:J02635; NID:G202591; PID:AAA40636.1; PID:G202592  
 J.Hayashida, K.; Okubo, H.; Noguchi, M.; Yoshida, H.; Kangawa, K.; Matsuo, H.; Sakaki, Y  
 J. Biol. Chem. 260, 14224-14229, 1985  
 A:Title: Molecular cloning of DNA complementary to rat alpha-2-macroglobulin mRNA.  
 A:Reference number: A05278; MUID:86033908; PMID:2414291  
 A:Accession: A05278  
 A:Molecule type: mRNA  
 A:Residues: 178-227/420-489, 'L', 491-526 <HA>  
 R.Kunz, D.; Zimmermann, R.; Heisig, M.; Heinrich, P.C.  
 Nucleic Acids Res. 17, 1121-1138, 1989  
 A:Title: Identification of the promoter sequences involved in the interleukin-6 dependen  
 A:Reference number: S03431; MUID:89160243; PMID:2466233  
 A:Accession: S03431  
 A:Molecule type: DNA  
 A:Residues: 1-22, 'V', 24-164 <XUN>  
 A:Cross-references: EMBL:X13983  
 A:Notes: the authors translated the codon GTG for residue 23 as Leu  
 R.Watburton, M.J.; Coles, B.; Dundas, S.R.; Gusterson, B.A.; O'Hare, M.J.  
 Eur. J. Biochem. 214, 803-809, 1993  
 A:Title: Hydrocortisone induces the synthesis of alpha(2)-macroglobulin by rat mammary m







FEBS Lett. 262, 349-352, 1990  
 A:Title: A genetic polymorphism in a functional domain of human pregnancy zone protein:  
 n.

A:Reference number: S09106; MUID:90242963; PMID:1692292

A:Accession: S09106

A:Molecule type: DNA

A:Residues: 668-690, 'M', 692-753 <MAR>

A:Cross-references: EMBL:X51541

R:Scitrip-Jensen, L.; Sand, O.; Kristensen, L.; Fey, G.H.

J. Biol. Chem. 264, 15781-15789, 1989

A:Title: The alpha-macroglobulin bait region. Sequence diversity and localization of c1e

A:Reference number: A34230; MUID:89380162; PMID:2476433

A:Accession: A34230

A:Molecule type: Protein

A:Residues: 670-752, 'Q', 754-759 <SOT>

C:Genetics:

A:Gene: GDB:PZP

A:Cross-references: GDB:120330; OMIM:176420

A:Map position: 12p13-12p12.2

A:introns: 73/3; 116/2; 146/3; 169/3; 703/1; 753/2

A>Note: the list of introns may be incomplete

C:Superfamily: alpha-2-macroglobulin

F:685-735/Region: bait region

Query Match 19.0%; Score 1396.5; DB 2; Length 1482;  
 Best local similarity 28.3%; Pred. No. 6.5e-71;  
 Matches 432; Conservative 280; Mismatches 587; Indels 225; Gaps 53;

12 LCCCTALAAVAPGRFVLTAPGIIIRPGANTTIGVLEHCPQVYTKAELTKASNLTV 71

14 LLLLSASDNSSTEQYVWLVPSLHTEAPKK-GCVLLSHNETVYVASLESGRERS- 71

72 SVLEAGVFEKGSFK--TLTLPPLNSADEIYELVGRTODELFSNTRLSFETKRI 129

72 --LFTDLVAEKDLFRICVSFTLPRISASSEVAFSLQIKGPTOD--FRKRVVLVLTQSL 127

130 SVFIQTDLALYKPKQEVKFRIVTLFSDPKPKTSNLTL-IDPKSNLLOQLSOQSDLVG 188

128 -VFQTDKPMYKPGQTVAFRVASVDENFRPNELPLILENPRNRJAQMSKLEAGI 186

189 ISKTFQLSHPIGLDMSIQOVNDQTYQ-SFOVSEYVLPFEVTLQTPLYCSMNSKHLN 247

187 NQLSFPLSSEPIQGSYRVVQTESGRIQHPFVEEFLPKFEKVVQPKLIISIMDEKVN 246

248 GTIRAKYVYKGVKDVLTFLPLSPFMGKNKIKTKFKINSANPFSNDEMKVMSNSN 307

247 IIVCGEYVYKGVKPELAT--VSLCKLSRVLNCQKQVCEEPS--QDL-----NSN 293

308 GL-----SEYLDLSSPG--PVELLTVTESVTGI--SHNVS--TNVFKQHDYIIEF 353

294 GCITQOVHTKMLQITNNGFEKMLRVEARIREGTDLEVTANRISBITNIVSK-----LKF 348

354 FDIYTVLKPNSLNFATVYTRADGNQLTLERRNNVITVTRQNTYTESGSSNG----- 408

349 VKVSHFQGGPIFFA--QVLLVDGKGVIPNKL--FPISVDANVYYSNATNNEGLOFS 404

409 -NQMKAAYQXI--NYTVPGSGTEKIEPILLEDSELOLKAVFISGSSMAVHSLFKSPS 464

405 INTTISISVNLKLVKRVFTVHPNLCFHYSW-VAEDHOGAO-----HNAKRVFSL-----S 451

465 KTYVQLKTRDENIKV-----SPEELVVSQNKELKELS--YMVVSRQGLVAVG----- 510

452 GSYHLEBVAGTLPQGHETITTAHTLNROMGSELSELSFHYLMAGVIVRSOTHTLPV 511

511 ----KQNSTMFSLPENSMTPRACVYVYIEDGELISDLKIPQVLVFNKKIKLYWSKV 566

512 ESGDMKGSFALSPVESDVADIAEMFIFAILPDGCVVGDSEKFEIENCLANKVDLSFSPA 571

567 KAEPSEKSLISVTPQPSIYGIYAVVDKSVNLMAANDITMENYVHEE----- 615

572 QSPASHAHLQVAA-PSGLCALRAVDQSVLLMKPEALSVSSVYNLLTVKDLTNPDPNV 630

616 -----LVNTGYLLGMFNMS---FAVFGCGMLVLTDAVLTQDYIDGVYDN 657

Db 631 DOOEEOGHCPRPFFIHNCAIYVPLSNBADIYSLFKMGLKFTNSKIRKPKSCVIPS 690

Qy 658 AE-----YAEFMEENG--HIVDI-HDFSIGSSGP-VRKHPPTMIWDT 699

Db 691 VSAQAVGGYAGGLGVERPYPQIGTVIVTFLNMQSSGPVETRVSYFETWIMELV 750

Qy 700 NMGRYIQEPEVTPDSITSMVATGVISDELGLTTPVLEQAFOPFIFLNPYSVI 759

Db 751 AVNSGVAEGVAVPPITTEWKAGAFCLSDALGISST-ASLRAPQFFVELTWPYSVI 809

Qy 760 RGEPALEITFNYLKDATEVKYIIEKSKFDILMTSSSEINATGHQCTLVSEDEATVL 819

Db 810 RGEVFTLKATVNLVLPKICIRVSVQLKASPAF-----LASQNTGSESSYCCGSE-ROTLIS 863

Qy 820 FPIRPTHGEPITVLAISPTA-----SDAVMTLVKAEIGEKYSQS 863

Db 864 WTVPTKLGNNVSVSHEAQSLQEGNEVEVEPEIKKDTVTKTLVLEAGIEQKITS 923

Qy 864 ILDLTDNRLQSTLKTLSFSPNVTYGSERVQITAGDVLGPSINGLASLIMPYGCGE 923

Db 924 ---SMTCASGANSBQSLSLKPSNVVKESARASFSVLGDLISGAMQNIQLMPGCGE 980

Qy 924 QNMVNFAPNIYIDYLTCKKQLTDNLKELASFMKQGYQRELLYQREDSFSAFG-NYDP 982

Db 981 QNMVNFAPNIYVLYNINETOQLTOEIKAKAVGYLITGYQRLNKKHQDSYSTFGERYGR 1040

Qy 983 S-GSTMASAFVLCPIEADPYIDIDONVLRHTYTMKGHQKSGEMDPGRVHSELOGG 1041

Db 1041 NQNTMTLAFVLTFTFQARSYIFIDEAHITQSLTWLSQMKDGCRRSGSLNNAIKGG 1100

Qy 1042 NKSPVTLTAYVTSLGKRYQPNIDVESIHLES-----EFSRGISDNVTLALITVA 1095

Db 1101 VEBEATISAVITLALIEIPVTPNPIVRNALPCLSEAMVAKGCTHG-SIVTYKALLAVA 1159

Qy 1096 LSSVGP-KAKELANMLTWBAEQQMGQVSSSEKLSD-----SQPRLDIEVAAYAL 1149

Db 1160 FSLGKQNRREILNLDKAVEDMLVHERPQRKAPGHLYQTOASAEVEMTSYVL 1219

Qy 1150 LSHFLOFQSEGI-P-----IMRWLSRQNSISGFPASTODTVALKASEF--ALL 1197

Db 1220 LAVL---TAQPAFTSGDLTSATNIVKIMKQNAQGSSTQDTPVALLHALSRYGAATP 1275

Qy 1198 MTERNTNIQVTVYGPSSPSP-----LAVQPMVAVN-----ISANGFPAICQLNV 1242

Db 1276 TRTEKT-AQVTVGDSDQTFSTNFPQVNNMLLLQOILPELPGSYVITVYGERCVIQTSM 1334

Qy 1243 VYNKASGSSRRRSIQNDFAFDLVAVXENKDLN--HYDLNVCTSPSGPR--SGMA 1297

Db 1335 KYNILPE-----KEDSPFALKVQTVPQTCGDKHAHTSPQISLTISYTGMRPASNMV 1385

Qy 1298 LMEVNLISGFMVSEALSL--SETVKVYEDHGKLNLYLDSVNEFOFCNIPAVANPKY 1354

Db 1386 IVDVKNVSGFTPLKFTVKMLERSSVSREVSNNHVLIVYEQVLTQTLSPSENVLDIDIV 1445

Qy 1355 SNTQDASVSIYDVYEPBRQAVRSY 1378

Db 1446 GDLKPAIVKYYDYETDESVAEY 1469

RESULT 12

A41185

alpha-2 macroglobulin M0G1 - mouse

C:Species: Mus musculus (house mouse)

C:Date: 28-May-1992 #sequence\_rev1stion 28-May-1992 #text change 12-Apr-1995

C:Accession: A41185

R:Overberg, L.; Torrekens, S.; Van Leuven, F.; Van den Bergh, H.

J. Biol. Chem. 266, 16903-16910, 1991

A:Title: Molecular characterization of the murinoglobulins.

A:Reference number: A41185; MUID:91358495; PMID:1840592

A:Accession: A41185

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1476 <OVR>  
A:Cross-references: GB:M65736  
C:Superfamily: alpha-2-macroglobulin

Query Match 19.0%; Score 1395.5; DB 2; Length 1476;  
Best Local Similarity 27.7%; Pred. No. 7.4e-71;  
Matches 434; Conservative 269; Mismatches 555; Indels 311; Gaps 50;

```

OY 13 LCVCTAALA-----VAPGRFLYAPGIIIRPGNVTIGVELLEHCPQVTKALLKT 65
DB 9 LCLFSLVLAFLHSAASLNDGSKMVLVPS-----QLYTTPKICHLHQLQANE 56
OY 66 ASNLIVSVLEABG-----VPEKSGFKTLT--LPSPLSADLEYELVGTGRQDEL 115
DB 57 TTVTVTASVLSQSGRKNLFDLVLVDKOLFQCVSEFIIIRLSSDEDEFLVDIKQPTHE--- 113
OY 116 FSNSTRLSFETKRISVFIOTDKALYKPKQKQKFRIVTLEFSDKPKYTSINT--LIDPKSN 174
DB 114 FSKRAVALVKNKESVVFQTDKPVYKPGQSVKFRVVMKMLRPLNELPLVYIEDPKKN 173
OY 175 LIQOMLSQSDLGVLISKTQFOLSHPILGMSIQV--QVNDQTYQSFOVEEVLPRKPEVT 233
DB 174 RIMQMRDITENGCLKMSFSLAEPICQFYKIVAKHSEKKEHSTVMEFVLPRFNVUL 233
OY 234 QTPPLYCSMNSKHLNGTITAKYTYGKPEKQDVLTFLPLSPWCKKNITTFKINGSANFS 293
DB 234 KVPNAMSVDDEVLSYACGKTYGKVPGHV-----KINVCERE 273
OY 294 FNDEKKNVMSNGIS---EYLDLSPGP-----VEILTTVESVTGI--SRNVSTN 341
DB 274 TQCREVNSQLD--NNGCSTQEVNITELQSKKNVEVOLFFHNATVIEEGGLSEFSRSGTYK 332
OY 342 VFFKQHDYIIEFDYTYVAKPSLNFATYKVRADGNOLTLERRANNVITVYQRYVEY 401
DB 333 IERTYKLI--FLKADSHRHGIPFVYKRLVDIKDPIPNK---VFYKAQELSYTA 386
OY 402 WSGNSNGCKMAVOKINYTPQSGTFKIEFPLLEDSELOKA-----YFLGS 450
DB 387 TTTDQHG-----LAFESIDTCISGS--LHKVNHKEHSCSYFCMBE 429
OY 451 KSMNAVHSLFK--SPKTYIOLKTRDENT-----KVQSPPELVVSGNKR--LKEI--SY 498
DB 430 RHASAGHVAAYVAYSLSKSYIYLDTERSSILPCNQIHTVQAHF--LIKGLDGLVKELEIFY 487
OY 499 MIVSREGVLAVGKQNSTM-----FSLTPNSMTPKACVIVVYIEDDELISDV 546
DB 488 LVMAHLSTIQTGHTHQVEBPGEAPVKGKALETPVEFSVPMKMYLITLIDGVIADS 547
OY 547 LKIPVQLVEKNKIKLYWSKVKAPESEKVALISVTQPSIVGIVAVDKSVNLMAKSNIDT 606
DB 548 VAPFEIEKCLRNVDLRFSTQSLPASQTRLQYVAS--PQSLCGLRAYDQSVLLKPESELS 606
OY 607 MENVAHELELYNTGYLLGMPMNSFAVFOCGMLVLTANILKYDIYDGVYDAAEYAE--- 662
DB 607 PS-----WYIYLP--GMQONKF-----VPSRLSEDOEDCIIYSSWIAEKHTN 647
OY 663 -----RPMEN-----BEHYIDH-----DESLGSSPH----- 685
DB 648 LVPHGTEKQVYRVEDMGLTAFNTNMIKLPIICFDGMVPIAPRVEFLAFTPEISWML 707
OY 686 -----VRKHPETMILDTNMGRYIAQVEEVTVPDSITSM 720
DB 708 RITLSRPEPRPKPDSSNDPLTERTRKTFPEETWVMDIYVNSGTGLAEVMTVPDITTM 767
OY 721 VATGVISDGLGLTTTPVELOAFQPFPIPLNLPSYVIRGEFALEITIFNYLKDATEV 780
DB 768 KAGALCISNDTGIGLSSV--VPLQAFKPFVEVSLPYSVVRGEAFMLKATVMYVILPFSQM 826
OY 781 KYIIEKSDKFDLMTSSEINATGHQOTLVPSBDGATVLPPIRPHLIGELIPTVYALST 840
DB 827 SVQLEASPDFTAVPGDD-----QDSYCLISANGRTSSMLVYPKSLGAVNVSVALEAQ 880
OY 841 AS-----DAVTQMILVKAEGIEKYSQSILDLTNRLOSTLTKLTSFSG 884

```

```

DB 881 SSEPQSEVATVPATGRKDTVVKVLYVPEIKQEHFTSSIFCASDAEIS---EXMSSGP 937
OY 885 PENTVGSERVQITAGIVLGPISINGLASLIRMPYGCGEOMINPAPITVILDTYTKKO 944
DB 938 PPTVVKSAHAHAFSFWGDLISALINTONPLHMPYGCGEOMNVLPAPITVYKLTNFTQQ 997
OY 945 LTDNLKERALSFMRQGYRELLYQREDSPSFAFG--NNDPQSGTWLSAFVRCPEADPY 1002
DB 998 LTQKIKTALGLRAGYQRELNTYKHKQDSYAFQDQNERGNTWLTFLVLSFAQARAF 1057
OY 1003 IDIDQNLHRTYTWLKHQKNGEFPDGRVHSELOCGKNSPTVLAIVYTSILGYRKY 1062
DB 1058 IFIDESHITHFTWLSQKQKNGCFRSGSGLFNNAKMGVDDEMTLSAIVTALLSSSLP 1117
OY 1063 QPNIDVQSHIFLES-----FESRGISQNTYLTALITYLSSVSGSP--KAKEALNMTWMAE 1116
DB 1118 ATHPVASKALSCLESQWKTIEOERNASVYVYKALMAVAFALAGQNKDEILKSLDEBAT 1177
OY 1117 QEGGMQFNV--SSESKLSDSW-----QPRSLDIEVAAYALLSHFLQFOTSEGIP----- 1163
DB 1178 KENNIIHKRPQKSKSEHHLKYQASAEYEMNAVYVLAAL-----TQAPAPSPEDLTLS 1233
OY 1164 --IKMWLSRQNSLQGFASDTDTVVALKALSEFAL--NMTERTNIQVTVTGPSSP--- 1215
DB 1234 MSTIMWLTQKQNSNGCFSTQDTVVALDALSKYGAVTFSRQKTLVTIQSTGFSQKQ 1293
OY 1216 ---SPLAVQPMAY-----NISANGRFALICQNLVYVYVYKASGSSRRRSIQNEAFD 1265
DB 1294 VENSRRLLQOVVALPDIPGDYTISSVSGCYVAAQTMRLYNNHLR-----KOLSAFA 1344
OY 1266 LDVAVK-----ENKQDLNHDVNLVCTSPFSGP--GRSGMALMEYNLISGFVWPSBAISLSSTV 1320
DB 1345 IWQTVPLTCNNPKKHSFQSLSTYSGSPASAMVYADVKMLSGF-----IPLKTV 1398
OY 1321 KKV-----EXDHGKLNLYLDSVNETQFCVNIPAVRNFKVNSTODASVSYVDYBPR 1371
DB 1399 KKLRLRHVSRTEVSNVNLVYLDQVTWQTLAFSFIQODIPVNLQPAIVKYVDYVETD 1458
OY 1372 ROAVNSYNS 1380
DB 1459 EWAFAYSS 1467

```

# RESULT 13

A20872  
ovostatin precursor - chicken  
N:Alternate names: ovomacroglobulin  
C:Species: Gallus gallus (chicken)  
C:Date: 05-Jun-1987 #sequence revision 13-Mar-1997 #text\_change 31-Mar-2000  
C:Accession: I50671; I50672; S29836; A20872; A33715  
R:Nielsen, K.L.; Sottcup-Jensen, L.; Nagase, H.; Thøgersen, H.C.; Ezerodt, M.  
DNA Seq. 5, 111-119, 1994  
A:Title: Amino acid sequence of hen ovomacroglobulin (ovostatin) deduced from cloned cDN  
A:Reference number: I50671; MUID:95218210; PMID:7535598  
A:Accession: I50671  
A>Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1473 <N1>  
A:Cross-references: EMBL:X78801; NID:G671863; PIDN:CAA5384.1; PID:G671864  
A:Accession: I50672  
A:Molecule type: mRNA  
A:Residues: 20-1473 <N1>  
A:Cross-references: EMBL:X78801; NID:G671863; PIDN:CAA5384.1; PID:G671865  
R:Nielsen, K.L.; Sottcup-Jensen, L.  
Biochim. Biophys. Acta 1162, 230-232, 1993  
A:Title: Evidence from sequence analysis that hen egg-white ovomacroglobulin (ovostatin)  
A:Reference number: S29836; MUID:93192299; PMID:7680377  
A:Accession: S29836  
A:Molecule type: Protein  
A:Residues: 977-980, 'N', 981-1007, 'M', 1009, 'K', 1010-1028 <N1>  
R:Nagase, H.; Harris Jr., E.D.; Woessner Jr., J.F.; Brew, K.

[illegible]

QY 683 SPVAKHPETWIM---LDTNMGYRIYOEREVVPQISITWVANGFVISEDGLGTTTP 739

Db 741 ---IKEFPETIMIDILLINSTG--KASVSTIIPDITETWKASAFVEBELAGCM-SVP 793

QY 740 VELQAFQEPFIFILNPSYVIRGEFALETITFNLYLKDATEVKYIIEKSKFDILMTSSEI 799

Db 794 ATLTAFOFPFVDTLTLPYSIIHGDFLVRANVFVYLNHCIKINVL-----L 839

QY 800 NATGHOQTLVISEDGA-----TVLFPIRPHLGEIPIITVTA-----L 837

Db 840 ESLDYOAKLISPEDDGCVCARIKRSYVWNIFFPKGTDDVLSIATETNDEACEEALRNI 899

QY 838 SPASDAVATQMLVYKAEIGEKYSOSILDLTGNRLQSTLKTISFPSPPTVTSERVQI 897

Db 900 RIDYDPTQIRALVEBEGIRRETONFLICMDVVIS--QDAVIDLPTVNEGSPRPSF 956

QY 898 TAIGDVLGPSINGLASLIRMPYCGCGEONMIFNAPNIYILDYLTTRKKQLTNLKEKALSF 957

Db 957 SVVGDMIGTALQNVHQLLQMPFGNGEQNVLFAPNIYVLDYLDKTRQLSBDVSKTIGYL 1016

QY 958 ROGYORELIIQREDGSSFAFGNYDPSGTWLSAFVLRCPLEADPYIDIDONVLRHTYTWL 1017

Db 1017 VSGYOKOLSKYKAPDGSYTFGIRDKGNWTWLTFAFYKSFSEASRFLYIDDNVQAOITLWL 1076

QY 1018 KGHQSNSEFMPQGRVHSELQGGNKSPTVLTLYITSLTGYKRYQONIDVQSIHLES 1077

Db 1077 ATKQKTDGCFQSTGLLVNNAMKGVENELSLSYITTLAEAGSHSHVIRNAFYCLET 1136

QY 1078 EFSRGISDNYTALTLYTVALSSVSPKAKEA-LNMLTWRAEOEGMOPWVSE-----SK 1130

Db 1137 ASEKNITDITDYQALVAVAFCLAGKABICSPFLRELQKSAKEVGDGSKWEGNQSRAPESH 1196

QY 1131 LSDSQPNRLDIEVAAYALLS-----HPIQFQTSSEGIPIYKMLSRQNSIGGFASTODTT 1185

Db 1197 LLDHYQ--STDVEITISYVLLALLYKENRQOEDLTTKASAIYQWIIROONSXGSGFASMDTV 1254

QY 1186 VALKRLSEFAL-----MNERITNIQV-TYTGPS-----SPESLAVVOPMAVNS 1229

Db 1255 VALQMLAAAGATYNSVTQNTVTKINSKNTPEKFTYNNENRLLLOQTPLPQV-PGKYSLT 1313

QY 1230 ANGFPFALCOLVVNVNVKASGSSRRRSIIONCEAFDLDVAVKEN-----KDDLNVHVLNV 1284

Db 1314 VNGTGCVLILQTLRLRNINHP-----ECAFQFSLSVQGRSNMASCPRDQPEKFIYV 1362

QY 1285 CTSPSGP-GRSGMALMEVNLISGFNVPSAISISE-----TVKKVEYDHGKLNLYDSV 1337

Db 1363 ISSYGRKSSSNWVIIDVYKMLSGF-VPKRS-SLDQLIDHTQWQVEKXKHVLLYYLGNIL 1420

QY 1338 ---NETQFCVNIIPAVRNPKVSNTOQASVSYDYVYPRRQAVKSYNSEVVLSSCDLGSVD 1393

Db 1421 QKRREVTFSVE---QDFVTVHHPKPAVOIYDYTTEBYAVAEWMS-----LCRGV 1468

RESULT 14

A42210

alpha-1-macroglobulin precursor - rat

C|Species: Rattus norvegicus (Norway rat)

C|Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #ext\_change 16-Jul-1999

C|Accession: A42210; A26124

R|Margesard, B.; Martin, N.; Johansson, S.

Biochemistry 31, 2346-2352, 1992

A|Title: cDNA cloning and sequencing of rat alpha-1-macroglobulin.

A|Reference number: A42210; MUID:92172859; PMID:1371696

A|Accession: A42210

A|Molecule type: mRNA

A|Residues: 1-1500 <VAR>

A|Cross-References: GB:M64000; GB:J05359; NID:g205383; PIDN:AAA1591.1; PID:g205384

A|Experimental source: strain Sprague-Dawley, liver

A|Note: sequence extracted from NCBI backbone (NCBI:87164)

R|Jonberg-Holm, K.; Reed, D.L.; Robert, R.C.; Hebert, R.R.; Hillman, M.C.; Kutney, R.M.

J. Biol. Chem. 262, 438-445, 1987

A|Title: Three high molecular weight protease inhibitors of rat plasma. Isolation, character

A|Reference number: A26124; MUID:87083487; PMID:2432067

A:Accession: A26124  
 A:Molecule type: protein  
 A:Residues: 25-29, 'F', 'J1-44 <LON>  
 C:Superfamily: alpha-2-macroglobulin  
 C:Keywords: glycoprotein; plasma; proteinase inhibitor  
 F:1-24/Domain: signal sequence #status predicted <SIG>  
 F:25-147/Product: alpha-1-macroglobulin #status predicted <MAT>

Query Match 18.8%; Score 1384; DB 2; Length 1500;  
 Best Local Similarity 27.8%; Pred. No. 3.4e-70;  
 Matches 434; Conservative 263; Mismatches 588; Indels 274; Gaps 50;

12 LVCCTAALAVAPGPRFLVTAPGIIIRPGGNVITIGVLEHSCSQT-VKAEILKTASNL 69  
 15 LILLPRDPAATAGKRRYVVLVPSSELYAGVEKVCVA-LNHLEMTVLANTLEYGVQYSNL 73  
 70 TVSVLEAGSEVFEKSGFKTLTLPSPLENSADEIYELRVGRTODEILFSTSLSPETKRI 129  
 74 LID---QAVDKOSSSCSSFTISRPL-SPSALIAVEIKGPTHFI---KKKSMITKAE 124  
 130 S-VFIQTDKALYKPKQEVFRIVTLFSDPKPYKTSNLT-IPKPSNLTQOMLSQOQSD- 186  
 125 SPVFQTDKPIKPKQYKFRVYVDISFRVNETFPVYIENPKNRIFQM-QVNDLP 182  
 187 -GVISKTQQLSSHPILGMSIQVQVND-QTYVQSFQVSEYVLPRKEVTLQTPLYCSMSK 244  
 183 GGLHQLSPFLSVEPALGIYKVVVQKDSGKKIHSFEVKEYVLPRKEVQVQMKPTMAFLBE 242  
 245 HANGITAKYTYGKPVKGVTL-----TFLPLSPFGK-KKNTITFKINGSANSPND- 296  
 243 ELVVTAAGLYTYGKPVGVLVTKVCRKTYQSYNSCHGQHSKICEFBSKQADEKCGFRQV 302  
 297 -----BEKKVMSNSGLSEYLDLSPPELITLVESVTGISRNS 339  
 303 VTKYVQPRQKDYDKIEVBAKIKEDGTG---IELTGSCELIANTISK-----LKFTK 353  
 340 TNVFKQHDYIIEFDYTVLK-----PSLNFATATKVRADGNQTLLEBRNNVITV 393  
 354 ANTFYRPG---LPFGQVLLVDEKQPIPNKQLTVQVNSVRSQFFTTDEHGLANILLDT 410  
 394 TORNTYVSGNSNGSNQKMEAVOKINITYTPQSGTFKIEPFLIEDSELOLAKYFSGSKS 453  
 411 T--NETFSFMG-----IRVYKQNNICFDMMWVDEYNT 441  
 454 MAVHSLFK--SPSKTYIOLKTRDENIKVSGPEELVY-----SGNRLKELS--YMWVS 503  
 442 QADHSAARITFSPRSRTIQLLEVLGTIACQTOEIRHFLHEDALDAADLTFFYILTKAR 501  
 504 GOLVAVGK-----QNSTMFSILTPENSWTPKACVIVYIIEDEGELISDVLKIPVL 553  
 502 GSIFNMSGHVPFLBEGKVKGVVSPFIRVBPMAFPAKILVITILLPRELIDAVOKFDEIK 561  
 554 VFKNKIKLYWSKVAKBPSEKSLRISVTQPSDIVIGIIVAVDKSVNIMNASNDITMENVAHE 613  
 562 CFANTVNLSPSAOSIPASDTHLTVMKAT-PLSLCALTAVDQSVLLLKPEAKLSPOSIVNL 620  
 614 L-ELVNTGYIYGMF-----MNSFAPQCGELMVLTD 643  
 621 LPQKABQAGVILGPLPYKGGENCICAEDITHNGIIVYTPKODLNDNAYSVFOSIGIKIFIN 680  
 644 ANLTQYIDGVYD-----NAEYARFMEENBGHIVDIDHPSLSS 683  
 661 TRVHPRCPMYQVAPPLPYGEPQALAMSAIPGAGY-----RSSNIRTSMMMMGAS 733  
 664 -----PHVRKHPETWILDTNMGRIYQEFEVTVPSDITSWATGVIISDLGIG 734  
 734 EVAQEVETVRYKTFPBTWMDVPLDLSGDGLPVKVPDITTEKSAFCLSGTGG 793  
 735 LTTTVELEDAQPPFFITLPLYSVIRGEFALEITFNILKATEKVKVILIESDKFDILM 794  
 794 LSTT--ISHKVPFPFLETLPLYSVIRGEFALEITFNILKATEKVKVILIESDKFDILM 852  
 795 TSSEINATGHQOTLLVPSSEDAATVLPFIRPTHLGELIPITVA---LSP----- 839

853 VGS-----HEDSHCIQNERKTVSMAYTPKSLGSEVNFATAEALQSPELCGNKVAEYVA 906  
 840 -TASDAVTOMILVKAEGIEKYSOSIILDTDNRLQSTLKTLSFSPGPNVTYGSRRVQIT 898  
 907 LVQKTVKVPYVIEBEGIEKECTYITLCPQDAEIQ---EMWTLDPANVEGSAKATQS 963  
 899 AIGDVLGPSINGLASLIMPYCGEONMINFAPNIYIIDLTYTKKQOTDNLKEKALSPMR 958  
 964 VLGDILGSAQNLQVLLQMPYCGEONNVLFPNIVYLEINETOQLTEALISKALSYLI 1023  
 999 QGYRELLIYQEBDSFSAFQ--YDPSGTSWLSAFVLRCELEADPYIDIDONVLRHYT 1015  
 1024 SGYQRLNVOHSDGSYSTFGGRKMSOGNTWLTFAVLKAPAOAQSYIIEKHTITNAFN 1083  
 1016 WLKHQKNGEFPWDGRVHSELOGKNSPTLTATYITSLIGYKQGNIDVOGSHFL 1075  
 1084 WLMKQRENGCFQOQSGSLNNAMKGVDEVTLSLITLLEMLPVTHSVVRNALFCL 1143  
 1076 ESEFSGISDN-----YTLALITVYLSVGS-PRKAEALNMLTWRAEOGGMQVWSES 1129  
 1144 ETAMA-SISNGBESHVTKALLAYAFALAGRARSEVLESINDAVNEESYHKQRPKN 1202  
 1130 -----KLSDSWQPR-SLDIEVAAYALLSHLOFOT-----SEGIPIMRWLS 1169  
 1203 VEENVRKMSFSPYKRPASAEVEMTAYVLLAYLTSASRPTRDLSSDLTASKIVKMS 1262  
 1170 RQNSLGGFASQODTVYALKALSEFALMNTERTNIQTYVYGPSSPSPLAVQPMANV-- 1227  
 1263 KOONSHGFPSTQDTVVALQALSKYGAATFT-KSNKEVSVTIESGT---VSGTLHAVNG 1318  
 1228 -----ISANGFQALCOLNVVNVAVASGSSRRRRTIONOEAFLDY- 1268  
 1319 NRLLQERLADLPBNYITTKYSGSCVYLQTSIKTNILPEA-----EGEAPFTLKVN 1370  
 1269 AVKENKDDLNH---VDLNVCTSPSGP-GRSGALMEVNLISGF--MVS-EAISLSEYK 1321  
 1371 TLPNFDKAEHRKQIHINYSYIGERPNSNMVIVDVGVSGFIPVKSVKLLQOQSNIQ 1430  
 1322 KYEYHGLNLVLSVNETQCVNIPAVNRKVSNTQDASISYVYIYEPROAVSNYS 1380  
 1431 RLEVNTNHLVLYIEKLTQTMGFSFAVEQDIPVKNLKBPVKVYVETDEPAIEYSVA 1489

RESULT 15  
 JC5144  
 murinoglobulin precursor - guinea pig  
 C:Species: Cavia porcellus (guinea pig)  
 C:Date: 31-Jan-1997 #sequence\_revision 27-Feb-1997 #text\_change 17-Mar-1999  
 R:Itawaki, H.; Suzuki, Y.; Sinochata, H.  
 J. Biochem. 120, 1167-1175, 1996  
 A>Title: Cloning and sequencing of cDNAs encoding plasma alpha-macroglobulin and murinog  
 A:Reference number: JC5143; PMID:97164019; PMID:9010766  
 A:Contents: liver  
 A:Accession: JC5144  
 A>Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-1464 <IMA>  
 A:Cross-references: DDBJ:D84339  
 C:Superfamily: alpha-2-macroglobulin  
 F:1-23/Domain: signal sequence #status predicted <SIG>

Query Match 18.7%; Score 1374.5; DB 2; Length 1464;  
 Best Local Similarity 27.9%; Pred. No. 1.1e-69;  
 Matches 430; Conservative 243; Mismatches 596; Indels 271; Gaps 49;  
 13 LVCCTAALAVAPGPRFLVTAPGIIIRPGGNVITIGVLEHSCSQT-VKAEILKTASNLVVS 72  
 16 LALLMTAASVHKQPVWVLP-----TQLYTEVEBKSCGLHLHYLKEVTVSAS 63  
 73 VLEA-----EGVEKSGFKTLTLPSPLENSADEIYELRVGRTODEILFSGNS 119

```

Db 64 LMSSMGKSLFSDPEVEHDLFCQVSVF---TLPTSSPNEVAFLSVQIKGRTH---TFSEE 117
Qy 120 TRLSFETKRSVFIQDTKALYKPKQOEKFRIVTLFSPFKP--YKTSILMILKDPKSNLIQ 178
Db 118 MAVVVRNTESTILIYQTDKPMKPGQYKFRVSVDRMLRQHHELFPLVYIEDPRKNIMQ 177
Qy 179 WLSQOSDLGVISKTPQLSSHPILCDWSIQV-QVNDQTYOSFOVSEYVLPKFEVTLQTP 237
Db 178 WRDITSENGIKQLSFNLSEBPIQGFYKIMVLKKSKEKEHFSAEYVLPFEVQVCPK 237
Qy 238 YCSNNSKHLNCTINAKTYGKPVKGDVTLTFLPLSPMGKKKNIKTETKINSANFSPN-- 295
Db 238 AVTLDEKRVNTVCGCKTYGKPVLGHV-----KNTCHLFKSYFGSSYGHNYG 285
Qy 296 ---DEEMKNVMSDNGLSSEYL-----DLSPGVEILTYTESVTGIS----- 335
Db 286 GQDCEBESQOLNQGCTTQAVRTSRDPMNMNHHINLQYNALITBEGTLEFGTGT 345
Qy 336 --RNVSTNVFPHDIIIEFPDYTVLKPSLNFATYKVTBAGNOULTERRNNVITV 393
Db 346 EVRKPRNTLVAVSTD-----SHFRPGIPFVQVQFLBDVKG---VPVADKHIFLKV 392
Qy 394 TORRYTERSGNSGNGKMEAVOKINTVPOSGTFKI-----EPLLEDSSELQ 443
Db 393 FTTKYN--SAYTDEHGLVEFSTNTTDLVAPSLFKYKYEKGHCFFKHYCIEELKQVDF 450
Qy 444 KAYFLGSKSMAVSHLSPKSPKTYIOLKTRDENIKVSGPFLVSV---GNKRLKLS-- 497
Db 451 VFY-----SVF-SYKSFVYLEPVTGALPCGQMTVHVHYTLNGVLOELRBMV 498
Qy 498 --YVWVSRQOLVAVGKONSTW-----FSLTPNSWTPKACVIVYIEDDEIISD 545
Db 499 FYULIMARGIIVQGTNAFMEPGELKGFNLISIPBSYMAPOMLIYAILPGEVIAD 558
Qy 546 VLKIPVQLVFNKIKLYWSKAKABSEKVSIRISVTOPDSIVGIVAVDKSNLNNASNDI 605
Db 559 SAKFEIENCLLSOVGLSFRPQSLPASQTHLRVT-APPQSLCALRAVDQSVLIMRPEAL 617
Qy 606 TMENVVHELEL-----YNTGYLGMFMN--SFAVFOEGCLW 639
Db 618 SPSSITVNLISIKDJDSSDLMLANFPENCQDINYDKIKSSLRNDKDTYKFIESMGLN 677
Qy 640 VLT-----DANLTKDYIDGVYDNA--EYAEFMEBNEGHIVDIHPSLGSSPHVRKHP 691
Db 678 IFTMLKTRDPLCPD--DLFPQVVAIKYTD-----HTVPABVDEKPLSETVRYTYP 727
Qy 692 ETWIMLDMNMGRIYQEFVTVPSITSWATGVVISDGLGLTTTPVELQAFQPPPIF 751
Db 728 ETWIMDLVEVNVSSGVAEVAVTPDITTEMKAGALCLPBDTGLGSP-ASLRAPQPFVE 786
Qy 752 LNLPSYVIRGEEFALETIFNYLKDATKVKYIIEKSDKFDILMTSEINATGHQOTLLVP 811
Db 787 LTMPSYVVRGEEVTFPKATVNLNYSKCIQVSVQLEASPAFTAVPLAKD-----QDSYCL 839
Qy 812 SEDG-ATVLPRIPTHLEIPITVYALSPAS-----DAVTQMLVKA 854
Db 840 CEDGRQVSMVLTVPKALGNVNFVSATQOQSELCGEVAAPPEARKKDVTVKPLVEPE 899
Qy 855 GIEKSYSOSILLDLTNRLOSTLKLSPSPPNVTYGSERVQITAGDVLGSPINGLASL 914
Db 900 GIKKEXYMNISFLCASDAVIS--EKLSTLKLPPKLVYDSARAFPSVFGDILSSSIKNTQNL 956
Qy 915 IRMPYSGEOMNINPAPRIYILDYLTKKKOLTDMLKEKALSFMROGYORELLYQREDSF 974
Db 957 LQMPYGCCEQNMVLPAPRIYVLKYNLNETDLOETKSALIGYLSAGYQROLNRYKHPDSY 1016
Qy 975 SAFG--NYDPSGSTWLSAFVLRCEFLADPYIDIDQNVLHRTYTLKGHOKSGEFPWDBGR 1032
Db 1017 SAFGHQNRNGNQTMLTAFVKTLAGARAFIFIDETHSHAFTWLSQKQKNGCGRSSGT 1076
Qy 1033 VHSSELQGNKSPVTLNAYIVTSLI--GYRKYQPNID-----VOESIHFLESEFSGISD 1085
Db 1077 LFHNDLKGVDDEVTLASAVITIALLEMBLPAHTPVVSKALTCINLESLE-MAKEGSG--SY 1134

```

```

Qy 1086 NYTLALITYALSSVGS--PRAKELANMLTWRAEGQMQFVWSSESKLSDS---WQPR--S 1139
Db 1135 VYTKALLAFAFALAGNODKRKEITLKSLEBAVKEDNSIHMERPOKRRPREALLYQFOAPS 1194
Qy 1140 LDIEVAAYLALSHFLOPQISEGIP-----IMRWLSRQNSLGGFASSTODTTVALK 1189
Db 1195 AEVEMTSYLAHL-----TAQGTPTPEEMTSAMRIYNNITKQONSYGFSSTODTTVALH 1250
Qy 1190 ALSEPAL-----MTERNTIQTVTGPSSPPLAVQ-----PMAVNI 1228
Db 1251 ALSRYGATRGRTKALTALVKIQSSGTFSTKFOV-----DNSNLLLOQVSLPHIPEEYTI 1305
Qy 1229 SANGFPAICQLNVVNVKASGSSRRRRRSIQNGEAPFLDY-AVKENKDDLN-HVDLNVCT 1286
Db 1306 SVSGEQCVYQTALKTNVFLE-----KEKYAFALQVHTVPQTCDDPKAKHRFQISL 1356
Qy 1287 SFGGPR--SGMALMEVNLISGFMYPSEALISL---SEYKAKVEYDHGKLNLYDSVNET 1340
Db 1357 EVSYTGRPASNMALVNVKMSIGFSPDKSTVKMLERSDHSVSRTEVKNVNLIVVDQVNE 1416
Qy 1341 QFCVNI PAVNFKVSNTOASVSIVDYIEPRRQAVRSYNS 1380
Db 1417 TLNFSFDVLDVPEVRNLKPAFVAVYDIYETGEAPAFETSA 1456

```

Search completed: January 15, 2004, 18:14:44  
Job time : 39 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2004, 18:03:24 ; Search time 18 Seconds

(without alignments)  
3730.785 Million cell updates/sec

Title: US-10-020-095-4

Perfect score: 7348  
Sequence: 1 MGGPPLITRAHLICVCTAAL.....HSSVIFCFCKLYEMELML 1428

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1455	19.8	1477	1 A113 RAT	P14046 rattus norv
2	1440	19.6	1474	1 A2MG HUMAN	P01023 homo sapien
3	1412.5	19.2	1472	1 A2MG RAT	P06238 rattus norv
4	1406.5	19.1	1476	1 A2M1 MOUSE	P28665 mus musculu
5	1402	19.1	1451	1 A2M2 MOUSE	P28666 mus musculu
6	1396.5	19.0	1482	1 P2P HUMAN	P20742 homo sapien
7	1395	19.0	1473	1 OVOS CHICK	P20740 gallus gall
8	1373.5	18.7	1495	1 A2MG MOUSE	P01638 mus musculu
9	1026.5	14.0	1673	1 CO3 LAMJA	P00685 lampraja ja
10	991.5	13.5	1744	1 CO4 HUMAN	P01028 homo sapien
11	987	13.4	1738	1 CO4 MOUSE	P01029 mus musculu
12	984.5	13.4	1640	1 CO3 ONCMY	P98093 oncorhynch
13	973.5	13.2	1651	1 CO3 NAANA	P01833 naja naja
14	966	13.1	1663	1 CO3 RAT	P01026 rattus norv
15	956	13.0	1663	1 CO3 MOUSE	P01027 mus musculu
16	950	12.9	1737	1 CO4 RAT	P08649 rattus norv
17	934	12.7	1663	1 CO3 HUMAN	P01024 homo sapien
18	921	12.5	1676	1 CO5 HUMAN	P01031 homo sapien
19	905	12.3	1680	1 CO5 MOUSE	P06684 mus musculu
20	904	12.3	1666	1 CO3 CAPVO	P12387 cavia porce
21	857.5	11.7	1620	1 CO3 EPTBU	P98094 epatretrus
22	543	7.4	920	1 CO4 BOVIN	P01030 bos taurus
23	467.5	6.4	1536	1 Y964 THEMA	P98079 thermocoga
24	432.5	5.9	726	1 CO3 RABIT	P12247 oryctolagus
25	290.5	4.0	1892	1 Y833 RICCIN	P92nd6 rickettsia
26	283	3.9	1653	1 YFHM ECOLI	P08440 anabaena sp
27	263.5	3.6	1906	1 YFA0 ANASP	P08440 anabaena sp
28	227	3.1	1582	1 YU30 RALSO	P08440 anabaena s
29	187	2.5	1516	1 Y819 PSEAE	P08440 anabaena s
30	183	2.5	1534	1 YFAS EGO57	P08440 anabaena s
31	180.5	2.5	1720	1 FTSH CHUVU	P56369 chlorrella v
32	176	2.4	1534	1 YFAS ECOLI	P56369 chlorrella v
33	174.5	2.4	2366	1 TOXB CLODI	P18177 clostridium

34	173.5	2.4	1905	1 Y659 PASMU	Q9cmz1 pasteurella
35	166.5	2.3	2358	1 YEBJ ECOLI	P76347 escherichia
36	164.5	2.2	2004	1 YP73 YERPE	Q8zdj2 yersinia pe
37	161	2.2	1337	1 P152 YEAST	P19685 saccharomyc
38	161	2.2	1562	1 YMA1 YEAST	Q04781 saccharomyc
39	161	2.2	2660	1 YEBJ EGO57	Q8xv7 escherichia
40	161	2.2	3354	1 CADN MOUSE	Q99p14 mus musculu
41	161	2.2	4349	1 FAT2 HUMAN	Q9nyq8 homo sapien
42	159.5	2.2	1609	1 FIG2 YEAST	P25653 saccharomyc
43	159	2.2	1656	1 ATG8 YEAST	Q12674 saccharomyc
44	159	2.2	2201	1 TENA HUMAN	P24821 homo sapien
45	159	2.2	3174	1 CHAC HUMAN	Q96f17 homo sapien

## ALIGNMENTS

RESULT 1  
ID A113 RAT STANDARD; PRT, 1477 AA.  
AC P14046;  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Alpha-1-inhibitor III precursor.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RC TISSUE=Liver;  
RX MEDLINE=89153707; PubMed=2631216;  
RA Bradlak T.A., Northmann W., Hudson G.O., Shiels B.R., Gehring M.R.,  
Ray G.H.;  
RT "Sequence and acute phase regulation of rat alpha 1-inhibitor III  
messenger RNA."  
RL J. Biol. Chem. 263:3999-4012(1988).  
CC - FUNCTION: PROTEASE INHIBITOR WITH A WIDE SPECTRUM OF PROTEIN  
TARGETS, WHICH ATTACHES THROUGH ITS THIOLESTER FUNCTION.  
CC - SUBUNIT: Monomer.  
CC - MISCELLANEOUS: THIS INHIBITOR HAS REDUCED PLASMA CONCENTRATIONS IN  
ACUTE PHASE.  
CC - SIMILARITY: TO OTHER PROTEINS OF THE ALPHA-MACROGLOBULIN FAMILY,  
INCLUDING COMPLEMENT COMPONENTS C3, C4, AND C5.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC EMBL, J03552; AAA0628.1; -  
CC EMBL, M2358; AAA0629.1; -  
CC PIR: A29952; A29952.  
CC HSRP, P01023; 1BV8.  
DR GO; GO:001714; F:wide-spectrum protease inhibitor activity; NAS.  
DR InterPro; IPR002890; A2M\_N.  
DR InterPro; IPR001599; MacroglloblnA2.  
DR Pfam; PF00207; A2M; 1.  
DR Pfam; PF01835; A2M\_N; 1.  
DR PROSITE; PS00477; ALPHA\_2\_MACROGLOBULIN; 1.  
KW Serine protease inhibitor; Glycoprotein; Plasma; Bait region; Signal;  
KW Liver; Thioester bond.  
FT SIGNAL 1 24  
FT CHAIN 25 1477 ALPHA-1-INHIBITOR III.  
FT DOMAIN 601 750 BAIT REGION (APPROXIMATELY).  
FT DISULFID 48 86 BY SIMILARITY.  
FT DISULFID 251 295 BY SIMILARITY.  
FT DISULFID 269 283 BY SIMILARITY.  
FT DISULFID 468 563 BY SIMILARITY.

```

FT DISULFID 595 774 BY SIMILARITY.
FT DISULFID 643 678 BY SIMILARITY.
FT DISULFID 850 886 BY SIMILARITY.
FT DISULFID 924 1324 BY SIMILARITY.
FT DISULFID 1082 1130 BY SIMILARITY.
FT DISULFID 1355 1470 BY SIMILARITY.
FT CROSSLINK 975 978 IsoGlutamyl cysteine thioester (Cys-Gln).
FT CARBOHYD 55 55 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 247 247 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 321 321 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 393 393 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 508 508 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 750 750 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 777 777 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 872 872 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 994 994 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1143 1143 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1314 1314 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1427 1427 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 4 4 D -> N.
FT VARIANT 87 87 V -> L.
FT VARIANT 137 137 M -> V.
SQ SEQUENCE 1477 AA; 163773 MW; 4DC05367C8385D2B CRC64;

```

```

Query Match 19.8% Score 145; DB 1; Length 1477;
Best Local Similarity 29.4% Pred. No. 8.4e-72;
Matches 457; Conservative 237; Mismatches 578; Indels 284; Gaps 50;

```

```

QY 13 LCVCTAALVAPRPFVLTAPGIIIRPGN---VTIGVELLEHCPQOVTVAELIKTASN 68
DB 9 LCLFSALIAFLPFLPFSIL-----NNGSKYMLVPSQLYTERPEKICLHLVHINETVT 59
QY 69 LTVVLEAEG-----VFEGSKFK--LTLPBLPLNSADEIETARTGTQEBILPSN 118
DB 60 VTALISQGRGRLKFDLVLVDKDLFHCVSFTIPRLPSSSEESJLDINEGAKHK--FSE 116
QY 119 STRLSEFKRISVFIQTDKALYKPKQEVKFRIVTLFSPFKPYKTSINL-----IKDKS 173
DB 117 RRVLVKKKESVFPQTDKPKMYKQSGYKFRVNSMDKRLH-----LNLFLPLAIIEDPKM 172
QY 174 NLIQMISQSGDLGVISTKTPOLSHPIIGDWSI-----QVQVNDQTYQSFQVSYVLPEK 229
DB 173 NRIQMOMQVKTENGLKQSLFSAPIQGPYKIVILKQSGVKE--HSFTVMEFVLPRF 229
QY 230 EVTLOTPLCYCMSNGHLNGITIAKTYTKPKYKGDVTLTFLPLSPFKGKGNITKFKINGS 289
DB 230 GVDKVPVAISVYDEIINVTACATYTYGKVPVGHVKISL-----CHGN 272
QY 290 ANFSFNDEMGKQVDS---SNGLS-----EYLDLSPGPEILTTVTESVTG 333
DB 273 PTFSEETKSGCKEDSRLDNNGCSTQEVNITEFOLKENYLMQ--APHVAIVTTEBGTG 330
QY 334 ISRVSTVNFVKQDYIIIEFPDYTTVLKPSLNFATYVKTADGNQLTLEERRNNVYTV 393
DB 331 SEFGSGRIEVERFRNKFLPKADSHFRHGIPIFKVYKLVDIKDPPIPNQ-----VLIKA 386
QY 394 TORVTEWSSNGSNQKMEAVQKINTVPOSGTFKIEFLPLEBSSELQKAYFLGSKSS 453
DB 387 RDAGYTAATTDHGLAK-----FSIDTNGISDYS--LNIKVYH--KEES 427
QY 454 MAVHS-----LFKSPSKTYIQLKTR-----DENIKVSPREL--VVSNG 490
DB 428 SCHSSCTAEHNAEHTAIVAVYSLSKSYITLDBAGLPCNQLHTVOAHFILGQVLGV 487
QY 491 KRLKELSYMVSRGQLVAVGK-----QNSTFSLTPRNSWTPKACVIVYIED 538
DB 488 LQGVLFHVLVMAQSIIGTGNHTHQVGEESQVGNFLALPVEFSWVPAVKMLIYTLIP 547
QY 539 DGEIISDVLKIPVQVLFKNIKILYWSKYKAPSEKVSIRISVTPDQSVIGIVAVDKSVNL 598
DB 548 DGEVIADSVKQVBEKCLRNKVLHLSFSPQSLPASQTHMRVTAS--POSICGLRAVDQSVLL 606

```

```

QY 599 -----MNASNDITMENNVVHELELYNTGY--LGMFNNSFAVFOECGLMV 640
DB 607 QKPEAEISPLIYDLPEMQDSNFIASNDPFBEDCYLQPIAREDDVRYRYRETGLMA 666
QY 641 LTTDANL-TKQYINGVYNAEYA-----ERFNEEGHLYVDIHDFLSSP----- 684
DB 667 FTMKIKLPTCYCNDYDMVFLAVPAVALDSSTRGMYESL--PVAVKS--PLPQEPKRP 724
QY 685 -----HYRKHPPETMILDTNMGRYIQEPEVTPDSITSVMVAGFISEDLGLT 736
DB 725 PPKDPVETIRNTPPETIMDWLTVANSQVTELEMTVPDITTEKACALCLSDTGLGLS 784
QY 737 TTPVELQAPFPFLFNLPSYVIRGEFALETIFNYLKDATEVKVLIIEKSKFDILMTS 796
DB 785 SV-ASFAQFPFVELIMPYSVIRGEAFTLKATVNLVLPSSLPAVLLEASPDF----- 837
QY 797 SEINATGHQOTLAVPSDQATVLPPIRPTHGELPIITVTLSS-----PTAS----- 842
DB 838 TAVPVENNQDSYCLGANGRHTSSMLVTPKSLGWNFSVSAEASQSPQCGSEVATVETG 897
QY 843 --DAVTQMLVKAEGIEKYSOSILLDTNRLQSTLTSFSPPTVTGSERVQITAI 900
DB 898 RKDTVAVKVLIVBPGIKHEHTFSSLCASDAEIS--ETLSLLPLPTVVDOSARAHFSVM 954
QY 901 GDVLGSPINGIASLIRMPYCGEGQNMIFAPNTIYILDYLTKKQTLDTNLKEKALSPNRQ 960
DB 955 GDILSSAKNTQNLQMPYCGEGQNMVLFAPNTIYLYLTNETQQLTEKISKALGYLRAG 1014
QY 961 YQRELLYQREDSGSAFGNYD--PSGSTWLSAVLRCFLEADYIDIDQVNLHRTYWLK 1018
DB 1015 YQRELLYQREDSGSAFGNYD--PSGSTWLSAVLRCFLEADYIDIDQVNLHRTYWLK 1018
QY 1019 GHQKSGEPMFDPKRVHISELQGNKSPVLTATYIVTSLIGRYKQNPID--VOESIHFL 1076
DB 1075 KQKDSCEFSNSSLNNAKGGVDEITLSAITYALL--ESSLPDTPVVSKALSCLE 1132
QY 1077 SEFSR-----GISDNTYLAITYALSVS--PRAKEALNMLTVRABQEGMQPW----- 1124
DB 1133 SSMENIEQGNNGSPVYTKALMAVAFALAGQEKRENLKSLDEKALKEKDSIMHERQKP 1192
QY 1125 VSSEKLSDSWQPRSLDIEVAAYALLSHFLQFTSBCIP-----IMRLSROKNS 1174
DB 1193 TKSEGYLYTP-QASSAEVEMSAVYVLAFL--TAQAPAPEDIALSMGTIKMLTKQONS 1247
QY 1175 LGSPASTQDPTVLLKLSER--AALMTERTNIOVTYTGSPSPBLAVVQ----- 1222
DB 1248 YGGSSTQDPTVALLDLSKYGATFSKQKTPBSVTVQSSGFSQKFOVDSNRLLQOVS 1307
QY 1223 -----PMAVNISANGFGFALCOLNVVNVKASGSSRRRSIQONQAPDLVAVK--ENK 1274
DB 1308 LPYIPGNYTVSGBGVVNAQTTLRNVPLE-----KQPAFALKVQTVPLTCNNP 1358
QY 1275 DDLANVDLVNCTSPSGP--GRSGMALMEVNLISGFMPVSEALISSEYKYK-----E 1324
DB 1359 KGNQSFQISLEISYMGSRPASNNVIADVKKLSGF-----IPLKPYTKLERLGHVSRTE 1412
QY 1325 YDHGKLNLYDSVNEHQFCVNIYAVNRFKYSNQDASVSIYDYYEPRQAVRSYNS 1380
DB 1413 VTTNNVLLYLDQVNTQTLSSFTIIQODIPVKNQPAIVKVDYETDEVAFAEYSS 1468

```

```

RESULT 2
A2MG_HUMAN STANDARD; PRT; 1474 AA.
AC P01023; Q13677;
DT 21-JUL-1986 (Rel. 01, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Alpha-2-macroglobulin precursor (Alpha-2-M).
GN A2M.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```



OK NCBI\_TaxID=9606;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=85190481; PubMed=2581245;  
 RA Kan C.-C., Solomon E., Belt K.T., Chain A.C., Hiorns L.R., Fey G.H.;  
 RT "Nucleotide sequence of cDNA encoding human alpha 2-macroglobulin and  
 RT assignment of the chromosomal locus."  
 RT Proc. Natl. Acad. Sci. U.S.A. 82:2282-2286(1985).  
 RN (2)  
 RP SEQUENCE FROM N.A., AND VARIANT HIS-704.  
 RX MEDLINE=92246939; PubMed=1374237;  
 RA Mathijs G., Devriendt K., Cassiman J.-J., van den Berghe H.,  
 RA Marynen P.;  
 RT "Structure of the human alpha-2 macroglobulin gene and its promoter."  
 RT Biochem. Biophys. Res. Commun. 184:596-603(1992).  
 RN (3)  
 RP SEQUENCE OF 24-1474, AND CHARACTERIZATION.  
 RX MEDLINE=84239807; PubMed=6203908;  
 RA Sottrup-Jensen L., Stepanik T.M., Kristensen T., Wierzbicki D.M.,  
 RA Jones C.M., Loeblad P.B., Magnusson S., Petersen T.E.;  
 RT "Primary structure of human alpha 2-macroglobulin. V. The complete  
 RT structure."  
 RT J. Biol. Chem. 259:8318-8327(1984).  
 RN (4)  
 RP ERRATUM.  
 RA Sottrup-Jensen L., Stepanik T.M., Kristensen T., Wierzbicki D.M.,  
 RA Jones C.M., Loeblad P.B., Magnusson S., Petersen T.E.;  
 RT J. Biol. Chem. 260:6500-6500(1985).  
 RN (5)  
 RP SEQUENCE OF 832-1474 FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=85219061; PubMed=2408344;  
 RA Bell G.I., Ball L.B., Sanchez-Pescador R., Merryweather J.P.,  
 RA Scott J., Eddy R.L., Shows T.B.;  
 RT "Human alpha 2-macroglobulin gene is located on chromosome 12."  
 RT Somet. Cell Mol. Genet. 11:285-289(1985).  
 RN (6)  
 RP SEQUENCE OF 672-747.  
 RX MEDLINE=90242963; PubMed=1692292;  
 RA Marynen P., Devriendt K., van den Berghe H., Cassiman J.J.;  
 RT "A genetic polymorphism in a functional domain of human pregnancy  
 RT zone protein: the bait region. Genomic structure of the bait domains  
 RT of human pregnancy zone protein and alpha 2 macroglobulin."  
 RT FEBS Lett. 262:349-352(1990).  
 RN (7)  
 RP INHIBITORY SITE.  
 RX MEDLINE=84030513; PubMed=6195065;  
 RA Virca G.D., Salvesen G.S., Travis J.;  
 RT "Human neutrophil elastase and cathepsin G cleavage sites in the bait  
 RT region of alpha 2-macroglobulin. Proposed structural limits of the  
 RT bait region."  
 RT Hoppe-Seyler's Z. Physiol. Chem. 364:1297-1302(1983).  
 RN (8)  
 RP INHIBITORY SITE.  
 RX MEDLINE=81212827; PubMed=6165619;  
 RA Sottrup-Jensen L., Loeblad P.B., Stepanik T.M., Petersen T.E.,  
 RA Magnusson S., Joernvall H.;  
 RT "Primary structure of the 'bait' region for proteinases in alpha 2-  
 RT macroglobulin. Nature of the complex."  
 RT FEBS Lett. 127:167-173(1981).  
 RN (9)  
 RP INHIBITORY SITE.  
 RX MEDLINE=81255805; PubMed=6167263;  
 RA Hall P.K., Nelles L.P., Travis J., Roberts R.C.;  
 RT "Proteolytic cleavage sites on alpha 2-macroglobulin resulting in  
 RT proteinase binding are different for trypsin and Staphylococcus  
 RT aureus V-8 proteinase."  
 RT Biochem. Biophys. Res. Commun. 100:8-16(1981).  
 RN (10)  
 RP INHIBITORY SITE.  
 RX MEDLINE=82095610; PubMed=6172288;  
 RA Mortensen S.B., Sottrup-Jensen L., Hansen H.F., Petersen T.E.,  
 RA Magnusson S.;

RT "Primary and secondary cleavage sites in the bait region of alpha 2-  
 RT macroglobulin."  
 RT FEBS Lett. 135:295-300(1981).  
 RN (11)  
 RP STRUCTURE BY NMR OF 1337-1474.  
 RX MEDLINE=99081553; PubMed=9865955;  
 RA Huang W., Dolmer K., Liao X., Gettins P.G.W.;  
 RT "Localization of basic residues required for receptor binding to the  
 RT single alpha-helix of the receptor binding domain of human alpha-2-  
 RT macroglobulin."  
 RT Protein Sci. 7:2602-2612(1998).  
 RN (12)  
 RP VARIANT ILE-1000.  
 RX MEDLINE=91187639; PubMed=1707161;  
 RA Poller W., Faber J.-P., Olek K.;  
 RT "Sequence polymorphism in the human alpha2-macroglobulin (A2M) gene."  
 RT Nucleic Acids Res. 19:198-198(1991).  
 RN (13)  
 RP VARIANTS TYR-972 AND ILE-1000.  
 RX MEDLINE=92128897; PubMed=1370808;  
 RA Poller W., Faber J.-P., Klobbeck G., Olek K.;  
 RT "Cloning of the human alpha 2-macroglobulin gene and detection of  
 RT mutations in two functional domains: the bait region and the  
 RT thiolester site."  
 RT Hum. Genet. 88:313-319(1992).  
 CC -1- FUNCTION: IS ABLE TO INHIBIT ALL FOUR CLASSES OF PROTEINASES BY A  
 CC UNIQUE "TRAPPING" MECHANISM. THIS PROTEIN HAS A PEPTIDE STRETCH,  
 CC CALLED THE "BAIT REGION" WHICH CONTAINS SPECIFIC CLEAVAGE SITES  
 CC FOR DIFFERENT PROTEINASES. WHEN A PROTEINASE CLEAVES THE BAIT  
 CC REGION, A CONFORMATIONAL CHANGE IS INDUCED IN THE PROTEIN WHICH  
 CC TRAPS THE PROTEINASE. THE ENTRAPPED ENZYME REMAINS ACTIVE AGAINST  
 CC LOW MOLECULAR WEIGHT SUBSTRATES (ACTIVITY AGAINST HIGH MOLECULAR  
 CC WEIGHT SUBSTRATES IS GREATLY REDUCED). FOLLOWING CLEAVAGE IN THE  
 CC BAIT REGION A THIOLESTER BOND IS HYDROLYZED AND MEDiates THE  
 CC COVALENT BINDING OF THE PROTEIN TO THE PROTEINASE.  
 CC -1- SUBUNIT: HOMOTETRAMER, WHICH CONSISTS OF TWO PAIRS OF DISULFIDE-  
 CC LINKED CHAINS.  
 CC -1- TISSUE SPECIFICITY: Plasma.  
 CC -1- MISCELLANEOUS: CONTRARY TO THE RAT PROTEIN, WHICH IS AN ACUTE  
 CC PHASE PROTEIN, THIS PROTEIN IS ALWAYS PRESENT AT HIGH LEVELS IN  
 CC CIRCULATION.  
 CC -1- SIMILARITY: TO OTHER PROTEINS OF THE ALPHA-MACROGLOBULIN FAMILY,  
 CC INCLUDING COMPLEMENT COMPONENTS C3, C4, AND C5.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL, M11313; AAA51551.1; -;  
 CC EMBL, Z11711; CAAT7774.1; -;  
 CC EMBL, M36501; AAA51552.1; -;  
 CC EMBL, X68728; CAA48670.1; -;  
 CC EMBL, X68729; CAA48670.1; JOINED.  
 CC PIR; A94033; MAHU.  
 CC PDB; 1BV8; 30-SEP-98.  
 CC SWISS-2DPAGE; P01023; HUMAN.  
 CC Genew; HGNC:7; A2M.  
 CC MIM; 103950; -;  
 CC GO; GO:0016975; F:alpha-2 macroglobulin; NMS.  
 CC InterPro; IPR002890; A2M\_N.  
 CC InterPro; IPR001599; Macrogloblna2.  
 CC Pfam; PF00207; A2M; 1.  
 CC Pfam; PF01835; A2M\_N; 1.  
 CC PROSITE; PS00477; ALPHA\_2\_MACROGLOBULIN; 1.  
 CC Serine protease inhibitor; Glycoprotein; Plasma; Bait region; Signal;  
 CC 3D-structure; Polymorphism; Thioester bond.  
 CC SIGNAL 1 23  
 CC CHAIN 24 1474 ALPHA-2-MACROGLOBULIN.  
 CC FT SITE 693 694 TRANSGUTAMINASE CROSS-LINKING SITE

```

FT  DOMAIN      690      728      (POTENTIAL)
FT  SITE        704      709      BAIT REGION.
FT  SITE        719      723      INHIBITORY.
FT  SITE        730      735      INHIBITORY.
FT  DISULFID    48       86
FT  DISULFID    251      299
FT  DISULFID    269      287
FT  DISULFID    278      431
FT  DISULFID    470      470
FT  DISULFID    563      563
FT  DISULFID    595      771
FT  DISULFID    642      689
FT  DISULFID    821      849
FT  DISULFID    847      883
FT  DISULFID    921      1321
FT  DISULFID    1079     1127
FT  DISULFID    1352     1467
FT  CROSSLINK    972      975
FT  CARBOHYD     55       55
FT  CARBOHYD     70       70
FT  CARBOHYD    247      247
FT  CARBOHYD    396      396
FT  CARBOHYD    410      410
FT  CARBOHYD    869      869
FT  CARBOHYD    991      991
FT  CARBOHYD   1424     1424
FT  VARIANT      704
FT  VARIANT      972
FT  VARIANT      972
FT  VARIANT     1000     1000
FT  CONFLICT     63       63
FT  CONFLICT    563      563
FT  CONFLICT    1148     1148
FT  CONFLICT    1195     1195
FT  STRAND      1341     1347
FT  TURN        1352     1353
FT  HELIX       1355     1359
FT  STRAND      1360     1369
FT  STRAND      1379     1384
FT  STRAND      1390     1391

Query Match      19.6%; Score 1440; DB 1; Length 1474;
Best Local Similarity 29.1%; Pred. No. 5.6e-71;
Matches 451; Conservative 266; Mismatches 551; Indels 284; Gaps 59;

```

```

QY  12 LVCVTAALAVAPGRFLVTAAGIIRPGANTIGVLELHCPGQVTVKAEILKTASNLTV 71
DB  15 LVLVLPDASVSGKRFQVWVLPVSLHT-ETTEKGCVLISYINERVYVSASLESVGRKSL 73
QY  72 -SVLEAGG-VFEKGSFKTLTLPVLSPLNSADE--IYELRYTGRTODEILPSNSTRLSFET 126
DB  74 FTDLEAENDVHCAVAF-----AVPKSSNEEMVFLTVQVKGPTQ--FKKRTTVWKN 124
QY  127 KRISVFIOTDKALKKPGQEVKFRIVTLFSDPKFKTISNLIL-IDPKSNLIGQLSQQSD 185
DB  125 EDLSVFPQTDKSIKPKQGTIVFRVSMDEHNPINELPLVYIDDPKGNRLAQWQSFOL 184
QY  186 LGVISTFOLSSHPLIGDMSIOVQVND--QTYVOSFOVSEVILPKFEVTLQTPLYCSMNS 243
DB  185 GGLKQFSPFLSSSEFPQGSYKVVQKSGGRTEH-PTVEEFVLKFEVQVIVPKIITILE 243
QY  244 KHLNGITIAKTYKPKVKGDTLTLFLPL-----SFWGK 276
DB  244 EEMNVSVCGIYTYGKPVGVHTVASICRKYSDASDCHGSDQAFCEKFSGQLNSHGFYQ 303
QY  277 KKNITTKFKFKINSANFSNDEEMQVNDSSNGLESEYLDLSSPGVEVILT-----TVTESV 331
DB  304 VK--TKVFOLR-----KEYEMK-----LHTEAQIOEGTVELTGROSSEIIRTRI 347

```

```

QY  332 TGISRVSTNVFPEKQHDYIIIEFPDYTVLKPSLNFTATVAKTRADGNQLTEERRNNVVI 391
DB  348 TKLS-FVKVDSHRQ--IPFFG-----QVRLVDKGVPIP-----NKVI 384
QY  392 TVYQRYVTEWGSNGSNQMGMAVO-KINYT--VPQSGTFKIEFP-----ILED 438
DB  385 FT-RGENANY--SNMTTBHGLVQPSINTTNWGTSLTVRVVYKORSPOCYQWVSEEH 441
QY  439 SEIQAKYFLGSKSMAVHSLFSPKTYIQLKTRDENIVGS-----PEELVVSQNK--R 492
DB  442 EBAHNTAY-----LVSPSKSYVHLEPMHELRPGHTQYVQAHYILANGTLLG 489
QY  493 LKELS--YWVVSQGLVANG-----KONST--MFSLTPENSWPKACVIVYIEDDG 540
DB  490 LKLSFYLLIMAKGIVRGTHGLLVKQEDMKGFHSISIPVKSDIAPVRLIIVAVLPTG 549
QY  541 EIIISDLKIPVQVLPFKNKIKLYMSKYKABSEKVSRISTQPDSTVGIYAVDQKSVNLMN 600
DB  550 DVIQDSAKDVENCILANKVDLSFSPQSPLASHAHLRVTA--PQSVCALRAVDQSVLLMK 608
QY  601 ASNDITMENVHBL-ELYNTGY-----YLGMFMS-----FA 631
DB  609 PDELSASSVYNLLPEKDLTFPGPLANDODDEICIRHNVYINGITYTPVSTINEKDMYS 668
QY  632 VPQECGLMWLTDAANTKDYIDGYDNAE-----YAEFMEENEGHIVDIDFSL 680
DB  669 FLBDMGLKAFNTSKIRKPKCPQLQOYEMHPEGLRVGFESDVWVGARHLYVHEE--- 725
QY  681 GSSPH--VRKHPEPWTIMLDTMKGRIYOEPEVYTPDSTSVNATGFIYSELGIGLTT 737
DB  726 ---PHETVAKYPEPWTIMDLVNVNSAGVAEVVTPDITTEKKAFAFCLSEBAGIGISS 782
QY  738 TPVELAOFOPPEFLNLPYSVIRGEFALBITFIYLYKATEVYKIIIEKSDKFDIMTSS 797
DB  783 T-ASLAPQFPFELTMPISVIRGEAFTLKATYANLIPKICIRSVOLDSAPFLAVPEK 841
QY  798 E-----INATGHOQLLVSESDGATVLPPIRPHLGEIPITYTALS-----PTA 841
DB  842 EQAPHICIANGRQ-----TVSMAVTPKSLGNVNFVSAEALSESQELCGTEVPSV 890
QY  842 -----SDATYQMLLVAAEGIEKYSOSILLDLTDLNLOSTLKLTSFPNTYTGSERVQ 896
DB  891 PEHGRDVTYKPLVPEEGLEKETFNSSLCPGSGEVS--EELSLKPEVNVSEARAS 947
QY  897 ITAIGDVLGPSINGLASLIRMPYCGEOMNINAPINITYLIDYTKKKOQLDNLKELASFP 956
DB  948 VSVLIGLILSAMQNTQNLQMPYCGEOMNIVLPAPINIVUDYINETOQLTPREKSAIGY 1007
QY  957 MROGYORELLYOREDSFSAFG-NYDPS-GSTWLSAFVLRCLFLEADPYIDIDQNVLAHRTY 1014
DB  1008 LNTGYQRLNTYKHYDGSYSTFGERYGRNQGNMTLTAFLVKTFAQAVAFIIDEAHITQAL 1067
QY  1015 TWLKGHOKSNGEWDGVRVYHSELQGNKSPVTLTYITSLLGYKIQBNIDVQESIH 1074
DB  1068 IWTISOROKNGCFRRSGSLINNAIKGVDEVTLSAYITIALLEIPLTVTHPVRNALFC 1127
QY  1075 LES-----EFSKSIDNTYLLITVALSSVGS-PAKEXALMNLTRABOEGMQW--- 1124
DB  1128 LESAMKTAQBGDHG-SHYTTKALAYAFALAGQDRKKEVYLSLNEAYKKNDSVHEWP 1186
QY  1125 VSSSEKLSDSWQPR--SLDIEVAAYALLSHPLQFQ--TSEGIP-----IMRWLSRQRNSL 1175
DB  1187 QKPAKAVGHFHEQOASAEVEMTSYVLLA-YLQAQAPFSEDLTSAITVKNWITKQONQ 1245
QY  1176 GGFASITQDTTVALKALSEFALMTERINIQVTVGSSPSPLAVVQ----- 1222
DB  1246 GGFSSITQDTTVALHALSKYGAATFT-RTGKAQVITIQSSGTFSSKQOVNNRLLQOVS 1304
QY  1223 ----PMAVVISANGPFAICQALNVVNVYKASGSSRRRRSRIQONEAFDLDVAVK---END 1275
DB  1305 LPFLPGEYSMKVTGBCCVLYQSLKNTNI-----LPEKEEFPFALGVQTLDPOTCD 1353
QY  1276 DLN-HVDLNVCTSFSGPG--RSGMALMEVNILSGFMVSEAISSL--SETVAKVEYDHG 1328

```

```

DB      1354 EPRKHSFQISUSVSTGSRSSANNAIVVKNVSGIPKPTVKLERNNHNSRREVSSN 1413
OY      1329 KNTLYDSVNETQFCYNIPAVNFKVSTQDASVSIVDYEPBROQVRSYNS 1380
DB      1414 HVLIVLDKXSNQTLISFTVVLQDVPVRLDKPAIVKYVDIYEDERAIAYNA 1465

RESULT 3
A2MG RAT
ID      A2MG_RAT          STANDARD;      PRT; 1472 AA.
AC      P06238;
DT      01-JAN-1988 (Rel. 06, Created)
DT      01-JAN-1988 (Rel. 06, Last sequence update)
DT      15-SEP-2003 (Rel. 42, Last annotation update)
DE      Alpha-2-macroglobulin precursor (Alpha-2-M).
GN      A2M.
OS      Rattus norvegicus (Rat).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OY      NCBI_TaxID=10116;
RN      [1]
RP      TISSUE=Liver;
RC      MEDLINE=87083488; PubMed=2432068;
RA      Gehring M.R., Shiels B.R., Northern W., de Bruijn M.H.L.,
RT      Kan C.-C., Chain A.C., Noonan D.J., Fey G.H.;
RT      "Sequence of rat liver alpha 2-macroglobulin and acute phase control
RT      of its messenger RNA.";
RL      J. Biol. Chem. 262:446-454(1987).
RN      [2]
RP      SEQUENCE OF 1-164 FROM N.A.
RC      STRAIN=wislar; TISSUE=Liver;
RA      MEDLINE=89160243; PubMed=2466233;
RT      Kunz D., Zimmermann R., Heisig M., Heinrich P.C.;
RT      "Identification of the promoter sequences involved in the
RT      interleukin-6 dependent expression of the rat alpha 2-macroglobulin
RT      gene.";
RL      Nucleic Acids Res. 17:1121-1138(1989).
RN      [3]
RP      SEQUENCE OF 178-227 AND 420-526 FROM N.A.
RC      MEDLINE=86033908; PubMed=2414291;
RA      Hayashida K., Okubo H., Noguchi M., Yoshida H., Kangawa K., Matsuo H.,
RT      Sakaki Y.;
RT      "Molecular cloning of DNA complementary to rat alpha 2-macroglobulin
RT      mRNA.";
RL      J. Biol. Chem. 260:14224-14229(1985).
RN      [4]
RP      FUNCTION: IS ABLE TO INHIBIT ALL FOUR CLASSES OF PROTEINASES BY A
RC      UNIQUE "TRAPPING" MECHANISM. THIS PROTEIN HAS A PEPTIDE STRETCH,
RA      CALLED THE 'BAIT REGION' WHICH CONTAINS SPECIFIC CLEAVAGE SITES
RT      FOR DIFFERENT PROTEINASES. WHEN A PROTEINASE CLEAVES THE BAIT
RT      REGION, A CONFORMATIONAL CHANGE IS INDUCED IN THE PROTEIN WHICH
RT      TRAPS THE PROTEINASE. THE ENTRAPPED ENZYME REMAINS ACTIVE AGAINST
RT      LOW MOLECULAR WEIGHT SUBSTRATES (ACTIVITY AGAINST HIGH MOLECULAR
RT      WEIGHT SUBSTRATES IS GREATLY REDUCED). FOLLOWING CLEAVAGE IN THE
RT      BAIT REGION A THIOLESTER BOND IS HYDROLYZED AND MEDIATES THE
RT      COVALENT BINDING OF THE PROTEIN TO THE PROTEINASE.
RT      -1- SUBUNIT: HOMOTETRAMER, WHICH CONSISTS OF TWO PAIRS OF DISULFIDE-
RT      LINKED CHAINS.
RT      -1- TISSUE SPECIFICITY: Plasma.
RT      -1- INDUCTION: BY INFLAMMATORY STIMULUS THE LEVEL OF THIS PROTEIN
RT      FIRST INCREASES, THEN DECREASES AFTER A MAXIMUM.
RT      -1- SIMILARITY: TO OTHER PROTEINS OF THE ALPHA-MACROGLOBULIN FAMILY,
RT      INCLUDING COMPLEMENT COMPONENTS C3, C4, AND C5.
RN      [5]
RP      THIS SWISS-PROT ENTRY IS COPYRIGHTED. IT IS PRODUCED THROUGH A COLLABORATION
RC      BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION
RT      AT THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS
RT      USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY
RT      MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL
RT      ENTITIES REQUIRES A LICENSE AGREEMENT (SEE http://www.isb-sib.ch/
RT      OR SEND AN EMAIL TO license@isb-sib.ch).

```

```

DR      EMBL; J02635; AAA40636.1; -.
DR      EMBL; X13983; CA32164.1; JOINED.
DR      EMBL; X13984; CA32164.1; JOINED.
DR      EMBL; X13985; CA32164.1; JOINED.
DR      EMBL; M11792; AAA40637.1; -.
DR      EMBL; M11793; AAA40638.1; -.
DR      PIR; A26122; A26122.
DR      HSSP; P01023; 1BV8.
DR      GO; GO:0016975; F:alpha-2 macroglobulin; NMS.
DR      InterPro; IPR002890; A2M_N.
DR      InterPro; IPR001599; Macrogloblna2.
DR      Pfam; PF00207; A2M; 1.
DR      Pfam; PF01835; A2M_N; 1.
DR      PROSITE; PS00477; ALPHA_2 MACROGLOBULIN; 1.
KW      Serine protease inhibitor, Glycoprotein, Plasma, Bait region, Signal,
KW      Thioester bond.
FT      SIGNAL          1..27
FT      CHAIN           28..1472
FT      DOMAIN          620..750
FT      DISULFID        52..90
FT      DISULFID        254..302
FT      DISULFID        272..290
FT      DISULFID        281..434
FT      DISULFID        473..473
FT      DISULFID        598..769
FT      DISULFID        647..694
FT      DISULFID        819..847
FT      DISULFID        845..881
FT      DISULFID        919..1319
FT      DISULFID        1077..1125
FT      DISULFID        1350..1465
FT      CROSSLINK       970..973
FT      CARBOHYD        59..74
FT      CARBOHYD        75..74
FT      CARBOHYD        250..250
FT      CARBOHYD        339..339
FT      CARBOHYD        651..651
FT      CARBOHYD        772..772
FT      CARBOHYD        867..867
FT      CARBOHYD        989..989
FT      CARBOHYD        1364..1364
FT      CARBOHYD        1422..1422
FT      CARBOHYD        1426..1426
FT      CONFLICT        23..23
FT      CONFLICT        490..490
FT      CONFLICT        1472..1472
FT      SEQUENCE       1472 AA; 163701 MW; D08F9A9AEB10FE0B CRC64;

Query Match          19.2%; Score 1412.5; DB 1; Length 1472;
Best Local Similarity 27.6%; Pred. No. 1.8e-69;
Matches 426; Conservative 275; Mismatches 580; Indels 265; Gaps 51.

OY      5 PLTNALHLCYCTALAAVAPGPRFLVTPAGIIRRGANVTIGVELHEHCPSOVTVKAEILK 64
DB      13 PLTNALHLCYCTALAAVAPGPRFLVTPAGIIRRGANVTIGVELHEHCPSOVTVKAEILK 70
OY      65 TASN--LTVSVLAEAGVEKESFKTLTPSLPLNSADSI--YELRYNGRTODELFNSST 120
DB      71 VHGNQSLTDLVVDLDLPHCTSF-----TVQSSDDELMFTYOVKATHE--FRROS 121
OY      121 RLSPETKRISVFIQTDKALVYKQEVKFRITVLSDFPKYKTSNLIILIKDPSNLIIOOW 179
DB      122 TVLVAKKESIVFAQDKRIYKPGQVVRVRSVLDSEFRLBELPLVLIOGPKNNRIQOW 181
OY      180 LSGQSDLVISKTPOLSHPILGWSIOVQVNDQTYQSFQVSEYVLPKREVTLOTPLY 238
DB      182 QNFNEGKIGKQSLSPFLSSSEFPQGSYKIVIRTESGRVHPSPVEFVLPKREVTLPBET 241
OY      239 CSMSKHLNGITLTKYTGKPVKGVNLTFL-----PLSPFGKKKNI----- 280
DB      242 ITILEENNVASCGIYTGKVPGRVTVNCRKYSNPNCFGESVAFCEKLSQQLDGRG 301
OY      281 -----TKTFKINSANFSFNDEEMKRVMSNGLSFYLDLSSGFPVEIITTVTESYTG 333

```

```

Db 302 CFSQLVKTSFOLK-RQSEYEMQOLDVHAKIOEGTGVEE---TGKG---LTKITRTITK 352
Qy 334 ISRWSTVTFKQHDYIEFPDYTTVLKPSLNFATATVTRADGNQLTLEBRANNVITV 393
Db 353 LS-----FVNDSHFROGIPFGQVILVVGKRPID----- 383
Qy 394 TORNTYEWSSGNSGNQMEAVQKINTVTPSGTFKLEFPLEED--SSELQKAYF-- 447
Db 384 ---YETIFIGAD-----FANLYINTTDDKGLAR--FSINTDDIMQSTLTVAKYKDS 431
Qy 448 -----LGSKSSMAVHSLFK--SPKTYIQKTRDKNIVKSPFEL---VSGN--K 491
Db 432 NACGFMWTEENWAMHTAYAVSPSPRSFLHESLPDKLQCDQTLQOAHYIINGEAMQ 491
Qy 492 RLKEL--SYWVVSQGVAVG-----KONSTM--FS--LTPENSWPKACVITYIEDD 539
Db 492 ELKELVFLYLMMAKGIYRAGTHVLPLKQGMGRHFSILLIMETDLAVVALVYALPN 551
Qy 540 GEISDVLIKIPVQVFNKKIKLYMSKYAEPSEKVSLSISVTPDPSIVGIADVKNLM 599
Db 552 GEVVDGTAKEYEINCLANKVDLFRPNAGLPATALLSV--MASPOSIGLRAVDQSVLM 610
Qy 600 NASDITMENV-----VHELELYNTGYILGEMFMS----- 629
Db 611 KPTELSSALTYDLLPVKDLTGPOGADQREEDTNGCVKQNDTYINGILYSPVQNTNEED 670
Qy 630 -FAVFOECGLWLTLDANLTKDYI--DGVYDNAEYARFEMEENEGHVIDHDSLSSPHVR 687
Db 671 MYGLKMGKLVFNINSIRKPKVCERLADNGIPRAHYILVSQSHDAFLSESSPTEITRR 730
Qy 688 KHPETWIMLDTNNGYRIYQEFVTPDPSITSWATGVFISDGLGLTTPVEILOAFQ 747
Db 731 SYFPEWIMDLVVDVAGVAEVEVTPPTITEMKAGACLSNDGLGLSPV--VQFOAFQ 789
Qy 748 FFILNLPYSYIRGEPELETTIRNYLKDATEVYVYIKSKPKFLIMSSINATGHO 807
Db 790 FEVELTMEYSYIRGEPELETTIRNYLKDATEVYVYIKSKPKFLIMSSINATGHO 844
Qy 808 LLVSEDAVLPFIRPFLHGEIPITVTALESPTA-----SDAVTOMILV 851
Db 845 -CICMNRHTASMAVIPSLSGNVFTVSAELNSKELCGNEVVPVPEQKQDITIKSILV 903
Qy 852 KAGSIEKYSQSILDLTDLNLOSTL-----KTLSPFPPTVTVGSEVOITAIADVLG 906
Db 904 EPEGLE-----EYTFNSILCPMGAEVELALALKPSDVVEBSAASATVGLDILGS 955
Qy 907 SINGLASIIRMPYCGEONMINPAPNIYILYLTKKQJLNLKXALSFROGYOELL 966
Db 956 AMQNTODLLKMPYCGEONMVLFPAPNIYILYLTDLNLOSTLTOEIKTKAIYANTGYORQLN 1015
Qy 967 YQREDSFSAFGNDPSG---STWLSAFVLRGFLADPYDIDONVLRHTYTWLKGHO 1022
Db 1016 YKRDGSGISAFG--DKPERNHANMTLTAFLVLSKSPAQAKKFIPIDEVHITQMLMLSQOQK 1073
Qy 1023 SNGEFMPDGRVYHSELQCGNKSPLYTLAYITVSLGYAKYOPNIDVQESIHFLSEF--- 1079
Db 1074 DNGCFRRSGSLNNAMKGVDEVTLSAYITILLESLSPTHTVVVNNALFCLDTAKMSA 1133
Qy 1080 ---SNGISDNTYLLALITLASSVSPKA--KEALNMLTWRAEOEGMOWFVSESLSLDS-- 1134
Db 1134 RGAGSGSHVYTKALLAYAFALAGQDQTEKELIKSIDEAIVEEDSVMTTRPQ--PFSVVG 1192
Qy 1135 --WQPR--SLDIEVAAYVALLSHPIQFOTSEGP-----IMRWLSRORNSLGSPAS 1180
Db 1193 LMYQOPATSAEVEMTAYVLLAYL---TTERAPQOEDITLTAAMLVKKMLTKQONSHGSS 1248
Qy 1181 TODTVALKALSER--AALMTERNTNIQVTVSPSSPFLAVVQ-----PM 1224
Db 1249 TODTVALKALSKGSAFTTAKKAQVYTISSGTFKPKGVNNNNQOLLQRYVTLPTVPG 1308
Qy 1225 AVNISANGFPAICQOLNVVNVKASGSSRRRSRIONQEPDLVAVK---ENKDLNV 1280
Db 1309 DYTVKVTGEGCVYIQTSLKYSV-----LPREEFPFAVVVOTLPTCTEDPRAHT 1357

```

```

Qy 1281 DLNACTSPSPGPG---RSGMALMEVNLISGFMVSEASISLSET---VKQYEDHDKNLTYL 1334
Db 1358 SFOISLINSISTYGRSRESNMAIADVKVSGFIPLKPTKMLERSVHVSRTFVSNHVLTYL 1417
Qy 1335 DSVNETOFVCNIPAVNRFKVSNQDASVSLVDYEPERRQAVRSYNS 1380
Db 1418 DKVSNQTVNLISFTVQODPIRDLKPAVVKVYDYERKDEFAVAKYSA 1463

RESULT 4
A2M1_MOUSE
ID A2M1_MOUSE STANDARD; PRT; 1476 AA.
AC P2665;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Muringlobulin 1 precursor (Mug1).
GN MUG1 OR MUG-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 28-57.
RC TISSUE=Liver;
RX MEDLINE=91358495; PubMed=1840592;
RA Overbergh L., Torrekens S., van Leuven F., van den Bergh H.;
RT "Molecular characterization of the muringlobulins.";
RL J. Biol. Chem. 266:16903-16910 (1991).
CC -1- FUNCTION: A PROTEINASE ACTIVATES THE INHIBITOR BY SPECIFIC
CC PROTEOLYSIS IN THE BAIT REGION, WHICH, BY AN UNKNOWN MECHANISM
CC LEADS TO REACTION AT THE CYSTEINYL-GLUTAMYL INTERNAL THIOLESTER
CC SITE AND TO A CONFORMATIONAL CHANGE, WHEREBY THE PROTEINASE IS
CC TRAPPED AND/OR COVALENTLY BOUND TO THE INHIBITOR. WHILE IN THE
CC TETRAMERIC PROTEINASE INHIBITORS STERIC INHIBITION IS SUFFICIENTLY
CC STRONG, MONOMERIC FORMS NEED A COVALENT LINKAGE BETWEEN THE
CC ACTIVATED GLUTAMYL RESIDUE OF THE ORIGINAL THIOLESTER AND A
CC TERMINAL AMINO GROUP OF A LYSINE OR ANOTHER NUCLEOPHILIC GROUP ON
CC THE PROTEINASE, FOR INHIBITION TO BE EFFECTIVE.
CC -1- SUBUNIT: Monomer.
CC -1- TISSUE SPECIFICITY: Plasma.
CC -1- SIMILARITY: TO OTHER PROTEINS OF THE ALPHA-MACROGLOBULIN FAMILY,
CC INCLUDING COMPLEMENT COMPONENTS C3, C4, AND C5.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M65736; AAA73048.1; -.
DR HGSP; P01023; IBV8.
DR MGD; MGI:99837; Mug1.
DR InterPro; IPR002890; A2M N.
DR InterPro; IPR001599; MacrogloblnA2.
DR Pfam; PF00207; A2M; 1.
DR Pfam; PF01835; A2M N; 1.
DR PROSITE; PS00477; ALPHA 2 MACROGLOBULIN; 1.
KW Serine protease inhibitor; Glycoprotein; Plasma; Bait region; Signal;
KW Multigene family; Thioester bond.
FT SIGNAL 1 27
FT CHAIN 28 1476 MURINGLOBULIN 1.
FT DOMAIN 677 734 BAIT REGION (APPROXIMATELY).
FT DISULFID 48 86 BY SIMILARITY.
FT DISULFID 251 276 BY SIMILARITY.
FT DISULFID 269 288 BY SIMILARITY.
FT DISULFID 461 555 BY SIMILARITY.
FT DISULFID 587 773 BY SIMILARITY.
FT DISULFID 634 680 BY SIMILARITY.
FT DISULFID 849 885 BY SIMILARITY.

```



RC TISSUE=Liver;  
 RA MEDLINE=91358495; PubMed=1840592;  
 RX Overbeigh L., Torrekens S., van Leuven F., van den Bergh H.;  
 RT "Molecular characterization of the murinoglobulin.";  
 RL J. Biol. Chem. 266:16903-16910(1991).  
 CC -1- FUNCTION: A PROTEINASE ACTIVATES THE INHIBITOR BY SPECIFIC  
 CC PROTEOLYSIS IN THE BAIT REGION, WHICH, BY AN UNKNOWN MECHANISM  
 CC LEADS TO REACTION AT THE CYSTEINYL-GLUTAMYL INTERNAL THIOL ESTER  
 CC SITE AND TO A CONFORMATIONAL CHANGE, WHEREBY THE PROTEINASE IS  
 CC TRAPPED AND/OR COVALENTLY BOUND TO THE INHIBITOR, WHILE IN THE  
 CC TETRAMERIC PROTEINASE INHIBITORS STERIC INHIBITION IS SUFFICIENTLY  
 CC STRONG, MONOMERIC FORMS NEED A COVALENT LINKAGE BETWEEN THE  
 CC ACTIVATED GLUTAMYL RESIDUE OF THE ORIGINAL THIOL ESTER AND A  
 CC TERMINAL AMINO GROUP OF A LYSINE OR ANOTHER NUCLEOPHILIC GROUP ON  
 CC THE PROTEINASE, FOR INHIBITION TO BE EFFECTIVE.  
 CC -1- SUBUNIT: Monomer.  
 CC -1- TISSUE SPECIFICITY: Plasma.  
 CC -1- SIMILARITY: TO OTHER PROTEINS OF THE ALPHA-MACROGLOBULIN FAMILY,  
 CC INCLUDING COMPLEMENT COMPONENTS C3, C4, AND C5.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/announcements/>  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC -----  
 CC EMBL; M65338; AA073041.1; -.  
 CC PIR; B41185; B41185.  
 CC HSSP; P01023; 1BV8.  
 CC MGD; MGI:99836; Muq2.  
 CC InterPro; IPR002890; A2M N.  
 CC InterPro; IPR001599; Macroglb1n2.  
 CC Pfam; PF00207; A2M, 1.  
 CC Pfam; PF01835; A2M\_N, 1.  
 CC PROSITE; PS00477; ALPHA\_2\_MACROGLOBULIN; 1.  
 CC KMW Serine protease inhibitor; Glycoprotein; Plasma; Bait region; signal;  
 CC Multigene family; Thioester bond.  
 FT SIGNAL 1 27 BY SIMILARITY.  
 FT CHAIN 28 1451 MURINOGLOBULIN 2.  
 FT DOMAIN 677 734 BAIT REGION (APPROXIMATELY).  
 FT DISULFID 48 86 BY SIMILARITY.  
 FT DISULFID 251 276 BY SIMILARITY.  
 FT DISULFID 269 288 BY SIMILARITY.  
 FT DISULFID 461 555 BY SIMILARITY.  
 FT DISULFID 587 773 BY SIMILARITY.  
 FT DISULFID 634 680 BY SIMILARITY.  
 FT DISULFID 849 885 BY SIMILARITY.  
 FT DISULFID 923 1274 BY SIMILARITY.  
 FT DISULFID 1081 1104 BY SIMILARITY.  
 FT DISULFID 1298 1444 BY SIMILARITY.  
 FT CROSSLINK 974 977 Isoglutamyl cysteine thioester (Cys-Gln)  
 FT (By similarity).  
 FT CARBOHYD 55 55 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 313 313 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 500 500 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 749 749 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 776 776 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 871 871 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1401 1401 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT SEQUENCE 1451 AA; 16336 MW; 9B2A82EFAA6258F6 CRC64;  
 SQ  
 Query Match 19.1%; Score 1402; DB 1; Length 1451;  
 Best Local Similarity 27.7%; Pred. No. 6.6e-69;  
 Matches 439; Conservative 250; Mismatches 527; Indels 370; Gaps 54;

Db 57 TVTVTASIVSQTQRRLFDLVDKDLFCVSTFIPLTNSPDEEBFLY-VDIGPHE-- 113  
 Qy 115 LFNSTRSLSPETRIISVFIOTDKALKYKPOEYFRITVLSDPKPKYTSUNI-LIDPKS 173  
 Db 114 -FSRRNANLVKAKESVVFQVDTDPYKPGQSVFRVNSMDKTRPLNELPLVIEDPK 172  
 Qy 174 NLIQWLQSQSDGLVISTKPOLSSHPILGDMSTIOVQ---VNDQTYOSFOVEEYVLPKF 229  
 Db 173 NRMQMRDITENGLKQMSLSLAEPIDGPKYLVHKQSVKKE---HSTVMEFVLPFR 229  
 Qy 230 EYTLQTPLYSNMSKHLNGTITLAKTYGKPKVDVTLTFLPLSPWKKKNTTFKNGS 289  
 Db 230 NVDLKYPNASTVNDVQLQVTCGKYTGKVPQGVNISICHFEAGKE-VNSKLDNNGC 288  
 Qy 290 ANSFNDEEMKNWMSNGLSVLDLSSPGEVILITLVESVYGISRN-----VST 340  
 Db 289 STQEVNITELQ-----SKRNEYVL---FHNATVIEEGTLEPNYGTTKIERITN 338  
 Qy 341 NVFFKQHDYIIEFDYTVLKPALNFATVKTTRADGNQLTERRNNVITVTRNYTE 400  
 Db 339 KLIFLKAD-----SHRRHGIPFFVKVRLVDIGDPINP-----RVPIKAQVLGYTS 385  
 Qy 401 YMSGNSGNQKMAVQKINTYVPQSGTFKIEPILDSSELQKAYFLGSKS----- 452  
 Db 386 TTTTQHGILAK-----FSIDTAGF-----SGSLIHKVNHKKDSCYFFYCM 428  
 Qy 453 -----SMAVHSLFKSPSKTYIQLKTRDENI-----KYGSPFELVSGNKR-LKE 495  
 Db 429 ERYASAEHVAVAYSL-----SKSYIYLVKETSILPCNOIHTVQAHF--ILKQDLGLVKE 482  
 Qy 496 L--SYMVSRGQLVAVGKQNTM-----FSLTPENSWTPKACVIVYIEDGE 541  
 Db 483 LVRYIYLVMAQSLIQGNHHTHOVEPGEAPYKGNFDEIPVEFPMAMAKMLITILPDGE 542  
 Qy 542 IISDVLIKIPVLQFNKKIKLYMSKVAEPSEKSLRISTVTPDSIVGIYAVDKSVNLMA 601  
 Db 543 VIDSVVFIEIKLRNVKVDLSFSSOSLPASQRLQVATAS--POSTGLRAVDOSVLLKP 601  
 Qy 602 SNDITMENNVHELNVNTGYILGMFNNSFAVPECGMLVTLNLRKDYIDGYDNAEVA 661  
 Db 602 EDELSPS-----WITNLP--GMQHNKF-----IPSSLSBREDCTLYSSWA 642  
 Qy 662 E-----RFMEENEGHIVDIHDFS----- 679  
 Db 643 EKHITDWPHREKQDVRVYED-----MDLKAFTNLKIKLPIKCFDASPMGPGKPDIAF 697  
 Qy 680 -----LGSSPH-----VRKHPETWIMLDTMGRITVQEPF 710  
 Db 698 SSEVSGTLQKSSKRPQPEEPREDPPPKDPLAETTRKYEPETWMDIVTVNSGVAEVE 757  
 Qy 711 VYVPDSTTSVAVAGVISEDIGLITTPVELQAPQPFIFLNLPSVIRGESEFALERTI 770  
 Db 758 MTPVDITTEKAKALCLSNDTGIGLSSV-VPLQAPQPFVEVSLPSVVRGAEFPMKAVV 816  
 Qy 771 FNYLKDAVEKVIIEKSDKFDIMTSSS-----INATG-HQQTLLVPSDEGATVLPPIRP 824  
 Db 817 MNTLIPMSRMSVQLEASPDPTAVPVGDHDHYSGLSANGHTSSWLV-----TP 864  
 Qy 825 THLGEIPITYVALSPAS-----DAVTQMIIVKAGIEKYSQSLTLDL 868  
 Db 865 KSLGNNVFSVSAQOSSSEPCGSEVATVPATGRKQTVVKKLVLEPBGIKQEHFNSLFC 924  
 Qy 869 TDNRLOSLTKLTSFSPPTNYTGSERVQITAGDVCPISNGLASLIRMPYGGEDQNMN 928  
 Db 925 SDAEIS---EKMSLVLPPTVVKOSARAHPSVMGDISSALKNQNLHMPYGGEDQNMV 981  
 Qy 929 FAPNIYIIDLITKKQQLTDNLKEKALSPMROGQRELLYOREDSFSFAFG--VYDPSGST 986  
 Db 982 FAPNIYIYKLTQDTQITQIKTKALGFLRAGIQRELYNTHKQGSYSABPDQNGREGNT 1041  
 Qy 987 WLSAPVLRCELEADPYIDIDQNVLAHRTYTWLKHQKSNSEFMWPGRVHSELOCGNKSPPV 1046





FT CONFLICT 691 691 V -> M (IN REF. 3).  
 FT CONFLICT 753 753 N -> Q (IN REF. 2).  
 SQ SEQUENCE 1482 AA; 163836 MW; 01773EF61D2FCC3B CRC64;  
 Query Match 19.0%; Score 1396.5; DB 1; Length 1482;  
 Best Local Similarity 28.3%; Pred. No. 1.4e-68;  
 Matches 432; Conservative 280; Mismatches 587; Indels 225; Gaps 53;  
 12 LCCCTALAAVAPRPLVTAGIIRPGCNVITVELLEHCPGSPVYTKAKELTKAASMLTV 71  
 14 LLLLSLSDSNSTEPQWLVLPVSLHTEAPKK-GCVLSHLENTVYASLESREKRS- 71  
 72 SVLEAGVFEKSGFR--TLPLSPLENSADEIYELRVTGRJODEILPNSNTRLSFETKRI 129  
 72 --LFTDLVAEKDLPHCVSFTLPRISSASEVAFSLQIKGPTD--FRKRVTLVLTNQSL 127  
 130 SVFIQTDKALYKPOEYKFRIVTLPSPDKPKYKTSNLT-1DQPKSNLQOQLSQOQDLGV 188  
 128 -VFQOTDKPMYKPGQVAFRVVSVSDENFRPNELIPLIYLENPRNRIRAOQSLKLAGI 186  
 189 ISKTFOLSHPIIDGMSIQOVNQTYQ-SFOYSEVYLPRFEVTLQTPLCMNSKHLN 247  
 187 NQLSPJSSSEPIQSYRVVQTESGRIQHPFTVEBEFLPKFEVAVQVPKTIISIMDEKVN 246  
 248 GTTAKTYTGKRVKGDVTLTFLPLSFMGKKKNTKFKINGSANFSFDEEKVNDSSN 307  
 247 ITVGEYTGKRPVGLAT--VSLCRKLSRVLCOKQCEPES--QQL-----NSN 293  
 308 GL-----SEYLDSSPG--PVELITTVESVTGI--SRNVG--TNVFPKQHDYIEF 353  
 294 GCITQOVHTKMLQITNCFEMKLRVEARIREGTDLEVTANRISIEITNIVSK-----LKF 348  
 354 FDTYTVLKPSLNFAVTAVTRADNQFLERRNNVTVYQRYATEVWSGNSG----- 408  
 349 VKVDSHFQGGIPFPA--QVLLVDGKVPINPKL--PFIADVANDYINATINNEGLOAFS 404  
 409 -NQKMAVOKI--NYTVPOSGTFKIEPPILEDSESELQKAVPLGSKSMAVHSLFKSPS 464  
 405 INTSISVNLKFAVAVFTYHPMLCFHYGW-VAEDHQGQ-----HTANRVFL-----S 451  
 465 KTYIOLKTRDENIVG-----SPELVVSGNKKLKEIS--YVAVSRQOLVAVG----- 510  
 452 GSYIHLEFVAGTLCFCHTEITTAHTYLRQAMGEISELSFHYLLMAGVIVRSCTHTLPV 511  
 511 ----KQNSTWESLPENSWPKACVIVYIEDDEIISDVLIKIVQVLFKKIKILVYSKV 566  
 512 ESGMKSPALSPVESDVAPIARMFPAIIPDEVDGDSKEFEIENCLANKVDLSFSPA 571  
 567 KAEPSEKSLRISTQPSIVIGIVAVDKSVNLMANASNDITMENVHELE----- 615  
 572 QSPFASHAHQVAAA-PQSLCALRAVQSVLMLKPEALSVSSVYNNLTVDLTPNDV 630  
 616 -----LYNTGYLLGMFMNS--FAVQECGLMVLTDANLTQYIDGYVDN 657  
 631 DQOEEOGHCHRPFIHNGAIYVPLSSNEADYIFLKMGMKLVFTNSKIRPKSCSVIPS 690  
 658 AE-----YAEFMENEG--HIVDI-HDSIGSSPH--VKHPEPMWIMLD 699  
 691 VSAGAVGGYVAGLGVVERPYVQDGLTYNVIPLNNESSGPVEVYASVPEPMWIMELV 750  
 700 NMGRIVQEFVETVDSITSWATGFIVSEDGLGLTTPPELOAFOPFFILMPYSVI 759  
 751 AVNSGVAEAVGVYVDTITTEKKAFAFCSIDAGISIST-ASLRAFQDFPELIMPISVI 809  
 760 RGEFPALEITIFNYLKATEVKVIEKSDKDFILMTSEINATGHQOTLVLVSEDGATVL 819  
 810 RGEVFTLKATVANTLPKICIRVQKASPAF-----LASQTKGESESYICGSE-RQTLIS 863  
 820 FPIRPHLGEIPITYALSPTA-----SDAVTQMIIVKAGIKESYSQS 863  
 864 WTVTPKTLGNVNFVSAAEQSLBELCNEVVEVEIRKRDVITLTLVAAGIIEQETFS 923  
 864 ILDLTLNRLQSTLKTLSFSPPTVTGSEVQITAIQDVLGSPINGLASLIRMPYCGE 923

Db 924 ---SMTASGANVSEQLSLPLPSNVVSESARASFVLDIGSAMONIQWLTMPYCGE 980  
 Qy 924 QNMNIPAPNIIYIDYLTKKQQLTDNLKEXALSPMGQYORELLYORBDGFSFAG-NYDP 982  
 Db 981 QNMVLPAPNIIYVLTNETQDLOEIKAKVGLITGYQOLYKQODGYSSTFGERYGR 1040  
 Qy 983 S-GSTMFAFVLCFLEADPYIDIDQNVLHRTYTWLKGHOKSGNCEFPDPRVHSELOGG 1041  
 Db 1041 NQGNWTMTAVLTFTFQARQYIFIDEAHITQSLTWLSQMKDGCRRSSGSLNMAIKGG 1100  
 Qy 1042 NKSPPVITAVITYVSLGKRYQPNIDVQSIHFLS-----EFSKGISDNTLALITTA 1095  
 Db 1101 VEDEATISAVYTTALLETIPLVNPIVRNALFCLSESMMVAKEGTGH-SHYVFKALLAYA 1159  
 Qy 1096 LSSVGP-SKAEALNMLTWABOEGGQFVVSSEKLSD-----SMQPRSLDEVALA 1149  
 Db 1160 PSLGKONQREILNSLDKAVREDNIVHERPQRPAPVGHLYOTQAPSAEVMSTYL 1219  
 Qy 1150 LSHFLQFQTSRGP-----IMRWLSRQNSLGGFASQDQTTVALKALSEF--AAL 1197  
 Db 1220 LAYL-----TQAPAPTSQDGLSATINIVKIMKQONAGGQFESTQDVTVALHALSRYGATF 1275  
 Qy 1198 MNTERTNIQVTVTGSPSP-----LAVQPMAN-----ISANGFPAICQALV 1242  
 Db 1276 TRTEKT-AQVTVQDSQTFSTNPFQVNNNILLLOQISLPELPGEVYITVGERCVYLTQSM 1334  
 Qy 1243 VYVVKASGSSRRRRSQQNGCAFQDLDVAVENKODLN--HYDLNVCSFSGSGR--SGMA 1297  
 Db 1335 KYVILPE-----KESFPALKQVTPQTCGHKAHRSFOISLTISTGRPASMNV 1385  
 Qy 1298 LMEVNLGFMVSEASL--SETVKAVEYDHGKLNLYDSVNETQFCVNIAPVNFV 1354  
 Db 1386 IVDVKNVSGRIPLKPIYKMLERSSSRTEVSNHNLIVYEQYTNQTLSSFNVLQDIPV 1445  
 Qy 1355 SNTQDASVSIQVYEPFRQAVRSY 1378  
 Db 1446 GDLKPAIVKYDYETDESVALEY 1469  
 RESULT 7  
 OVOS CHICK  
 ID OVOS CHICK STANDARD; PRT; 1473 AA.  
 AC P20740;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last annotation update)  
 DE Ovisatlin precursor (Ovomacroglobulin).  
 OC Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 ON NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Oviduct;  
 RX MEDLINE=95218210; PubMed=7535598;  
 RA Nielsen K.L., Sottrup-Jensen L., Nagase H., Thøgersen H.C.,  
 R Etzerodt M.,  
 RT "Amino acid sequence of hen ovomacroglobulin (ovostatin) deduced from  
 RL DNA Seq. 5:111-119(1994).  
 RN [2]  
 RP SEQUENCE OF 37-49.  
 RC TISSUE=Egg white; PubMed=6408074;  
 RX MEDLINE=8328315; PubMed=6408074;  
 RA Nagase H., Harris B.D. Jr., Woessner J.F., Brew K.,  
 RT "Ovostatin: a novel proteinase inhibitor from chicken egg white. I.  
 RT Purification, physicochemical properties, and tissue distribution of  
 RL ovostatin".  
 RN J. Biol. Chem. 258:7481-7489(1983).  
 RN [3]  
 RP SEQUENCE OF 976-1028.

CC TISSUE=Egg white; PubMed=7660577;  
 CC MEDLINE=93192299; Sottrup-Jensen L.;  
 CC "Evidence from sequence analysis that hen egg-white ovomacroglobulin  
 CC (ovostatin) is devoid of an internal beta-Cys-gamma-Glu thiol  
 CC ester." Biochem. Acta 1162:230-232(1993).  
 CC BL000101. Biophys. Acta 1162:230-232(1993).  
 CC -1- FUNCTION: IS ABLE TO INHIBIT ALL FOUR CLASSES OF PROTEINASES BY A  
 CC UNIQUE "TRAPPING" MECHANISM. THIS PROTEIN HAS A PEPTIDE STRETCH,  
 CC CALLED THE 'BAIT REGION' WHICH CONTAINS SPECIFIC CLEAVAGE SITES  
 CC FOR DIFFERENT PROTEINASES. WHEN A PROTEINASE CLEAVES THE BAIT  
 CC REGION, A CONFORMATIONAL CHANGE IS INDUCED IN THE PROTEIN WHICH  
 CC TRAPS THE PROTEINASE. THE ENTRAPPED ENZYME REMAINS ACTIVE AGAINST  
 CC LOW MOLECULAR WEIGHT SUBSTRATES (ACTIVITY AGAINST HIGH MOLECULAR  
 CC WEIGHT SUBSTRATES IS GREATLY REDUCED).  
 CC -1- SUBUNIT: HOMOTETRAMER, WHICH CONSISTS OF TWO PAIRS OF DISULFIDE-  
 CC LINKED CHAINS.  
 CC -1- PTM: THIS PROTEIN LACKS THE THIOESTER BOND FOUND IN OTHER MEMBERS  
 CC OF THIS FAMILY.  
 CC -1- PTM: CONTAINS 56 MOL GLUCOSAMINE PER MOL SUBUNIT.  
 CC -1- SIMILARITY: TO OTHER PROTEINS OF THE ALPHA-MACROGLOBULIN FAMILY,  
 CC INCLUDING COMPLEMENT COMPONENTS C3, C4, AND C5.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL, X78801; CAA55384.1; "  
 CC EMBL, X78801; CAA55385.1; ALT\_INIT.  
 CC PIR, I50671; A20872.  
 CC HSPB, P01023; I8V8.  
 CC GO, GO:001714; P-wide-spectrum protease inhibitor activity; NAS.  
 CC InterPro, IPR002890; A2M\_N.  
 CC InterPro, IPR001599; MacroglublnA2.  
 CC Pfam, PF00207; A2M; 1.  
 CC Pfam, PF01835; A2M\_N; 1.  
 CC PROSITE, PS00477; ALPHA\_2\_MACROGLOBULIN, FALSE NEG.  
 CC DR Serine protease inhibitor; Glycoprotein; Bait region; Signal.  
 CC KW SIGNAL 1 36  
 CC FT CHAIN 37 1473  
 CC FT CARBOHYD 67 67 OVOSTATIN.  
 CC FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 89 89 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 131 131 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 134 134 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 191 191 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 342 342 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 403 403 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 527 527 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 568 568 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 757 757 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 1141 1141 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 1221 1221 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 1315 1315 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 1347 1347 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CONFLICT 46 46 M -> V (IN REF. 2).  
 CC SQ SEQUENCE 1473 AA; 166354 MW; A33C6847A14179BF CRC64;

Query Match 19.0%; Score 1395; DB 1; Length 1473;  
 Best Local Similarity 27.0%; Pred. No. 1,66-68; Indels 308; Gaps 59;  
 Matches 421; Conservative 309; Mismatches 522;

Db 127 E--SF-----VFVQTDKPIYKPGOSVMEFRVVALDFNFKPQVEMYPFLAVDPQNNRI.FQ 188  
 179 WISQOSDLGVISKTQPSLSHPILGDMSTQV-QVNDQTYGOSQVSEVYLKREYTLQPL 237  
 189 WQNVTSINIVIEIEPLTEEPILGNKIIIVTKSGERTSHPLVEEYVLPKEDVITVTAQ 248  
 238 YSNMSKILNGITTKATYTGKPKGVKDVLTFL-PLSFNGKK-----NITKPKINGSA 290  
 249 SLTWDSLETIVKICAVITYYGQVEKQVLSVRDSDISGRCKSKSYCSFTDLDTGL 308  
 291 NFSFDE--EMKNVDSNGISEYLDLSSPGVEILITVTSVTSISRVNSTVFFKQHD 348  
 309 SHILSKVFEIURI-----GYRKNLD-----VKAIIVKEQVCLTATQGISI--TQVM 355  
 349 YIEFFDVTYVLAKEPLNTATYKVTADQNGQLTEERANNVITYTORNYTEYSGSNG 408  
 356 SSIQENVDHHRGIPFGQIKLVKDKNSP-----SNKYQLFVNKNKTH----- 402  
 409 NQKMAVQKINTVPSG--FKIEPILDSSELQKAYFLG-----KS 452  
 403 -----NFTDINGIAPSIDTSKIFD-PELSLALYKTSQCHSEGTIEPSYDPA 451  
 453 SNAVSLKPSPEKTYIOLKTRDENIKVSP-----FEIIVSGNRLKELS--YMWVSRG 504  
 452 SLISVGRIV-SWTSFVRIEPLMKDSCQKMITVYIILNTEGHEHINIVNYYGMKAG 510  
 505 QLVAVK-----QNST-MFSLTPENSWPKACVITYYIEDDEIISDVLTQVQLVF 555  
 511 KIVLGEIKVNIQADQNGTFMPLVNVNKKMAPALILVYVMPAKELVADSVRFSEIKCF 570  
 556 KKKIKLWYSKYAESEKVSRLISVTPDSIVGIYAVDQSVNLMAASNDITMENV--VHE 613  
 571 KKKVQLQSEKQMLTTSNVSLVIEAA-ANSPCAVAVDQSKMLLSSELSSETTYNLP 629  
 614 L-----ELVNTGYVL-----GMFNNSFAVFCGLMVLTD 643  
 630 IQDLQGYIFNGILNLEDDQDPCVSSDILFHNGLYRPLTSLGSPVYQFLRMGKKFFTN 689  
 644 ANL-----TKDYI-----DGYVNDAAEYAEERMEENEGHIVDHFSLGS 682  
 630 SKIROPTVCTRETPVPSYFLNAGFTASTHVKLSAEVARE--ERKSHILET----- 740  
 683 SHVAKHPEPTW---LDTNMGYIYQEFVTVSDSTTSWATGTVFVISEDGLGLTTP 739  
 741 --IEFPEPTWIMDIILINSTG---KASVSTIPDITTEWASAFCEELAGFGW-SVP 793  
 740 VELQAFQFFFLNLPYGVIRGEEFALETTFNYLKQATEVVIIEKSDKFDILMTSSBI 799  
 794 ATLTFQFPFVDLTPYRIIHGEDLVANAVNYLNHCIKINVL-----L 839  
 800 NATGQOQLLVSEDEGA-----TVLPRIPTHGELPIVTVTA-----L 837  
 840 ESLDVAQLISPEDGCGCAKIRKSYVNNIPKRGVDFLSITATNDDEACEBALNII 899  
 838 SPTADAVTQMLVABEIEKYSQSILDLTDNRLOSTKLKLSFFPNYTGSEEROI 897  
 900 RIDYVDYQIRLILVBEIRRETONFLICMKDVIS--DPAIDLPTNVVGGSPRSPF 956  
 898 TAIGDVLPSINGLASLIRMPYCGEOMINPAPNIIYLDYTKKQPLTDNKEKALSS 957  
 957 SVVDIMGTALIONVQLOMPFGNCEQNNVLPAPNIVYLDYDKTKRQSLSEBVGKTIQYL 1016  
 958 RQGYRELLLYOREDSFSAFGNYPDSGTWLSAFVLRGFLPADPYIDIDONVLRHTYWL 1017  
 1017 VSGYQKQSYKHPDQSYSTFGIRDEKGTWLTAFYFKSFAEASRTYIDDNVQAQTLWL 1076  
 1018 KGHQSNSEFMDPGVHISELOGKNSPVITATVITVLSLGRKQPNIDQESLHPLFS 1077  
 1077 ATKQKTDCCFQSTGILVNNANKGAVENELSLASVITTLALLEGHSMSTVIRNAPCLET 1136  
 1078 EFSRGISNNTYLTALTYVLSVSGSPKAEALNMLTWBADEGGQGFVWSSE-----SK 1130

```
Db      113/ ASSEKINDIDYIQTALVAFAFLGAGAEICESTBLRLOQSAEVDOSKXMEQNQRAPBESH 1196
Qy      1131 LSDSWQPRLSDIEVAAVALLS-----HFLQFQISEGIPIRMKULSRORNSUGGFASDTDTT 1185
Db      1197 LLDHVQ--STDFVEITSYVLTLALKYPKNRSQEDLTAKSAIVQMIRROONSXGGFSMSODTV 1254
Qy      1186 VALKAASEPAL-----MNTERTNIQV-TLVGPS-----SPSLPNAVVOPMANVIS 1229
Db      1255 VALDALAAYGAATYSTONVIKINSKTFEKVITNNENLLLIQQTFPLPOV-PGKYSLT 1313
Qy      1230 ANGEFAFIQLNVVVNNAKSASSRRBSRIQOEAFDIDLVAIVEN-----KDILNHVDLNV 1284
Db      1314 VNGVCGVILIQAALKNYHLR-----EGAFGSLSVSQGTNASCPHPQPKFPIYL 1362
Qy      1285 CTSFSGP-GRSGLMLMEVNILLSGMVDSIAISLE----TVKKVEYDHGKLNIYDSV- 1337
Db      1363 ISSVTGKRSSNMVIYIDVKMSGF-VPYKS-SLDQLIIDHTVMQVEYKHNLVLYLINIL 1420
Qy      1338 ----NEUFCVINAVRFKSNTPODASVSYDVYPERRAVRNRSYNSEVALLSCDLCSDY 1393
Db      1421 QKRREKVFYSVE---ODEVTHKPAPVQIYDYETEDEVAAVEMS-----LCRGV 1468

RESULT 8
A2MG_MOUSE STANDARD; PRT; 1495 AA.
AC Q61838; Q60628;
AD 01-NOV-1997 (Rel. 35, Created)
BT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Alpha-2-macroglobulin precursor (Alpha-2-M).
GN A2M.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 25-44 AND 1240-1259.
RC TISSUE=Liver;
RX MEDLINE=93076803; PubMed=1280217;
RA van Leeuwen F.; Torrekens S.; Overbergh L.; Lorent K.; de Strooper B.;
RA van den Berghe H.;
RT "The primary sequence and the subunit structure of mouse alpha-2-
RT macroglobulin, deduced from protein sequencing of the isolated
RT subunits and from molecular cloning of the cDNA.";
RL Eur. J. Biochem. 210:319-327(1992).
RN [2]
RP SEQUENCE OP 1-161 FROM N.A.
RC STRAIN=J29/J;
RX MEDLINE=95095249; PubMed=7528166;
RA Umans L.; Serneels L.; Hilliker C.; Stas L.; Overbergh L.;
RA de Strooper B.; van Leeuwen F.; van den Berghe H.;
RT "Molecular cloning of the mouse gene coding for alpha 2-macroglobulin
RT and targeting of the gene in embryonic stem cells.";
RL Genomics 22:519-529(1994).
CC -I- FUNCTION IS ABLE TO INHIBIT ALL FOUR CLASSES OF PROTEINASES BY A
UNIQUE "TRAPPING" MECHANISM. THIS PROTEIN HAS A PEPTIDE STRETCH,
CALLED THE 'BAIT REGION' WHICH CONTAINS SPECIFIC CLEAVAGE SITES
FOR DIFFERENT PROTEINASES. WHEN A PROTEINASE CLEAVES THE BAIT
REGION, A CONFORMATIONAL CHANGE IS INDUCED IN THE PROTEIN WHICH
TRAPS THE PROTEINASE. THE ENTRAPPED ENZYME REMAINS ACTIVE AGAINST
LOW MOLECULAR WEIGHT SUBSTRATES (ACTIVITY AGAINST HIGH MOLECULAR
WEIGHT SUBSTRATES IS GREATLY REDUCED). FOLLOWING CLEAVAGE IN THE
BAIT REGION A THIOLESTER BOND IS HYDROLYZED AND MEDIATES THE
COVALENT BINDING OF THE PROTEIN TO THE PROTEINASE.
CC -I- TISSUE SPECIFICITY: Plasma.
CC -I- DEVELOPMENTAL STAGE: CONTRARY TO THE RAT PROTEIN, WHICH IS AN
ACTIVE PHASE PROTEIN, THIS PROTEIN IS ALWAYS PRESENT AT HIGH LEVELS
IN CIRCULATION.
CC -I- SIMILARITY: TO OTHER PROTEINS OF THE ALPHA-MACROGLOBULIN FAMILY,
INCLUDING COMPLEMENT COMPONENTS C3, C4, AND C5.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
```

[illegible]

```

QY 243 SKHNTITAKTYGKPVKGDVTLTFLPSFGKKKNTKTKFINGSANFEND----EB 298
DB 241 EELPTTACGVYTKPVPGLVTL-----RVCKKYSRRYSTCHGNQMSICAR 288
QY 299 MKQVMDSSNGLSREYL-----DLSSPG---PVELITVTTE-----SVTGISRANSTVFPK 345
DB 289 FSGQADBKCCFQVYVTKVFOQSOKGHDKIEVEAKIKEBGIGIEITLGSCCEIANALSK 348
QY 346 QHDYIIIEPDYTTVTLKPSINFTATVKTTRADGNQLTEERRNNVTVTQRYNTEYSGS 405
DB 349 -----LKFVKVNTNRPGLPFGSQVLLVDEKGPFP-----NKKITSVSPBLGTLSTFTTD 399
QY 406 NSGNQKMEAVOKINTVTPQSGTFKIEFPLEDSSELQAKAYFGSGSMAVHS--LEKSP 463
DB 400 EHGHLANT-SIDTSNFTAP-----FLRVVTVYKONHVCYDNMMWIDEFHTQADHSATLVFSP 453
QY 464 SKTYIQOLKT-----RDENIKVSGPFELVYSGNRKLELSRMVSVRQQLAVGVQNST 515
DB 454 SOSYIQLELVFGTLACGQTOEIRIHLYLMEDIMKNKKTLLTFYLLIKARG---SIGMLGSH 510
QY 516 MFSLEPEN-----SWTPKACVIVYIEEDGELISDVLKIPQVLVFNKIKLY 562
DB 511 VLSLBGNMKGVFSLPIQVEPGMAPRAQLIYAILNEBELVADQNFELKCFANKVNS 570
QY 563 WSKVKAEPSEKVSRLISVTQPSIVGIVAVDKSVNL-----599
DB 571 FSAQSLPASPDTHLKAKAA-PLSLCALTVADQSVLLLKPEAKLPSQISVNLPGKTVQGA 629
QY 600 -----NASNDITMENNVHELNTYTGILGMFNASPAVPECGMLVLTDLANTLK--- 648
DB 630 FPGVPYKCHENCISGSDITHNGIIVTPKHSIG-DMDAHSIFQSVGINIFTMSKIKRPF 688
QY 649 -----DYIDGVYDAA-----EYARFMEENEGHIVDHDPS 679
DB 689 CQEPQHYPMGVAAPALAAVAAAGPSSSRAMGVPMAGLDYSER-----NOVREKRE-- 741
QY 680 IGSSPHVRKHPETWIM---LDTNMGYRIYQEFVTVDSITSWATGFSVISEDGLGI 735
DB 742 -----TVRKYPETWIMDLVPLDVSGD-----GELAVKVDPTITEMKASAFCSGTTGLDS 792
QY 736 TTTTPELOAFOPFTITLNPYSYIRGEERALETTITNYLKDATEVYIIEKSDKPIILMT 795
DB 793 SST-ISTLOAFOPFEITLTPYSVVRGEAFTLKATVNMVSHCQIIVDLISPDF----- 846
QY 796 SSEINATGHOOTLVSESDGATVLPPIRPHLGEIDITVATASPTA-----841
DB 847 -LAAPVGGHSHHCIGNERKTVSMVATPKSLGEVFTTAELBSQELCGNKULTVPAL 905
QY 842 --SDAVTOMILVKAEGIEKSYSSQSLDLTDNRLOSTLKLTSFSPFPNTVTSSEVOITA 899
DB 906 VHKDITVVKSVIIEPBGIEKQIYNTLLCPQDTLQD---NSSLELPPNVVSGSARATHAV 962
QY 900 TGDVIGPSINGLASLIRMPYGCGBOMINPAPITYILDYITKKQOLTDMLKEKALSFMRO 959
DB 963 LDDIIGASAMQNLQNLQMPYGCGBOMVLFVPNIYVLYNLTQQLTEAIKSKAINVLS 1022
QY 960 GYORELTYOREDSFSAFGNY---DSSGSTWLSAFLRCFLMDPYIDIDQNLHTHTYTM 1016
DB 1023 GYOROLANTYHSDGISTFGNHGGGNTPGNTWLAFLAKPAQOSHIFLEKHTITNAFWM 1082
QY 1017 LKGHQKNGEFMDPGRVHISELQGNKSPVTLTAVYVTSILGKRYQOPNDVQESIHFE 1076
DB 1083 LEMKQKENGCFQOSGVLNANNAMGVDDDEVTLISAVITIALLEHPLVTHSAVKNALFCE 1142
QY 1077 SEPSRGISN-----YTLALITYALSSVGS--PRAKALNMLTWRAQEGMGQFVWS--- 1126
DB 1143 TAMA-SISQOSQESHVYTKALLAVAFALAGKAKRSSELISINDKADAVEESLHWQPGDV 1201
QY 1127 SESKSDSMQPR--SIDIEVAALISHLQOPF-----SEGPIPMKWLSRON 1173
DB 1202 QKKAISFQOPRPSAEVENTAVVLLAVYLTSESSRPTRLDISSDLSTASKIVWISKQON 1261
QY 1174 SLGFASTQDVTVALKALSEFALNMTERNTIQVTVTGPS-----SPSPLAVVQ 1222

```

```

DB 1262 SOGGLLLTQDVTVALQALSKYGSATFT--RSQKEVLTSSSGTFSKTFHVNSGNRLLQOE 1320
QY 1223 PMAVNISAN-----GFGPAICOLNVNVNKKASGSSRRRSRISIO-NOEAPDLDAVAVEND 1275
DB 1321 VRLPDLPGNYVTKGSSGCVYIQTSLKYNILPVADKAPPAQVNTVLPFLFDA-----E 1375
QY 1276 DLNHDVLANCTSPSGP-GRSGNALMEVNLISGF--WVPS-EA1SLSEYKVEYDHQKLN 1331
DB 1376 DHRTFQIRINVSYTGERPSSNNVIVDVKNVSGFIPMKPSPVKRLQDPNIOQTEVNTNHYL 1435
QY 1332 LYLDVNETQFCVNIIPAVNPFVYSTODASVSIVDYIEPRQAVNSYNS 1380
DB 1436 IYIEKLNTQTLGFSFAVEQDIPVKLKPAPIKYDYETDEFTVEEYSA 1484

RESULT 9
C03_LAMJA
ID C03_LAMJA STANDARD; PRT; 1673 AA.
AC 000685;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Complement C3 precursor [Contains: C3a anaphylatoxin] (Fragment).
OS Lampetra japonica (Japanese lamprey) (Encephaleus japonicus).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Lethenteron.
OX NCBI_TaxID=94989;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=92251197; PubMed=1578150;
RA Nonaka M., Takahashi M.;
RT "Complete complementary DNA sequence of the third component of
RT complement of lamprey. Implication for the evolution of thioester
RT containing proteins."
RL J. Immunol. 148:3290-3295 (1992).
CC -I- FUNCTION: C3 PLAYS A CENTRAL ROLE IN THE ACTIVATION OF THE
CC COMPLEMENT SYSTEM. AFTER ACTIVATION (C3B), IT CAN BIND COVALENTLY,
CC VIA ITS REACTIVE THIOLESTER, TO CELL SURFACE CARBOHYDRATES OR
CC IMMUNE AGGREGATES. CYCLOSTOMATES C3 APPEARS TO REPRESENT THE
CC COMMON ANCESTOR OF MAMMALIAN C3 AND C4, SHOWING SIMILARITIES TO
CC BOTH PROTEINS.
CC -I- SIMILARITY: TO C4, C5 AND ALPHA-2-MACROGLOBULIN.
CC -I- SIMILARITY: Contains 1 anaphylatoxin-like domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
DB EMBL; D10687; BAA00983.1; -
DB PIR; I50806; I50806.
DB HSSP; P01024; IC3D.
DB InterPro; IPR002890; A2M_N.
DB InterPro; IPR000020; Anaphylatoxin.
DB InterPro; IPR001599; MacroglobinA2.
DB InterPro; IPR001134; Netrin_C.
DB Pfam; PF00207; A2M; 1.
DB Pfam; PF01835; A2M_N; 1.
DB Pfam; PF01821; ANATO; 1.
DB Pfam; PF01759; NTR; 1.
DB ProDom; PD003264; Anaphylatoxin; 1.
DB SMART; SM00104; ANATO; 1.
DB SMART; SM00643; C35C; 1.
DB PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
DB PROSITE; PS01177; ANAPHYLATOXIN_1; 1.
DB PROSITE; PS01178; ANAPHYLATOXIN_2; 1.
KW Complement pathway; Plasma; Inflammatory response; Glycoprotein;

```



RA Belt K.T., Carroll M.C., Porter R.R.;  
RT "The structural basis of the multiple forms of human complement  
RT component C4.";  
RL Cell 36:907-914(1984).  
RN [12]  
RP SEQUENCE FROM N.A. (C4A).  
RX MEDLINE=91108039; PubMed=1988494;  
RA Yu C.Y.;  
RT "The complete exon-intron structure of a human complement component  
RT C4A gene. DNA sequences, polymorphism, and linkage to the  
RT 21-hydroxylase gene.";  
RL J. Immunol. 146:1057-1066(1991).  
RN [3]  
RP SEQUENCE FROM N.A. (C4B).  
RX TISSUE=Blood;  
RC MEDLINE=96163032; PubMed=8575831;  
RA Ugielati D., Townsend D.C., Christiansen F.T., Dawkins R.L.,  
RA Abraham L.U.;  
RT "Complete sequence of the complement C4 gene from the HLA-A1, B8,  
RT C4A00, C4B1, DR3 haplotype.";  
RL Immunogenetics 43:250-252(1996).  
RN [4]  
RP SEQUENCE FROM N.A. (C4B).  
RX Rowen L., Dankers C., Baekin D., Faust J., Loretz C., Ahearn M.E.,  
RA Banta A., Swartzell S., Smith T.M., Spies T., Hood L.;  
RT "Sequence determination of 300 kilobases of the human class III MHC  
RT locus.";  
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A. (C4A).  
RX Barlow K.;  
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE OF 1-22 AND 1056-1225 FROM N.A.  
RX MEDLINE=85156269; PubMed=3838531; Porter R.R.;  
RA Belt K.T., Yu C.Y., Carroll M.C.;  
RT "Polymorphism of human complement component C4.";  
RL Immunogenetics 21:173-180(1985).  
RN [7]  
RP SEQUENCE OF 680-756.  
RX MEDLINE=81264286; PubMed=6167582;  
RA Moon K.E., Gorzski J.P., Hugli T.E.;  
RT "Complete primary structure of human C4a anaphylatoxin.";  
RL J. Biol. Chem. 256:8685-8692(1981).  
RN [8]  
RP SEQUENCE OF 957-1044.  
RX MEDLINE=82182029; PubMed=6978711; Porter R.R.;  
RA Campbell R.D., Gagnon J., Porter R.R.;  
RT "Amino acid sequence around the thiol and reactive acyl groups of  
RT human complement component C4.";  
RL Biochem. J. 199:359-370(1981).  
RN [9]  
RP SEQUENCE OF 990-1037.  
RX MEDLINE=82150875; PubMed=6950384; Harrison R.A., Thomas M.L., Tack B.F.;  
RT "Sequence determination of the cholesterol site of the fourth  
RT component of human complement.";  
RL Proc. Natl. Acad. Sci. U.S.A. 78:7388-7392(1981).  
RN [10]  
RP SEQUENCE OF 1-21 FROM N.A.  
RX MEDLINE=94282044; PubMed=8012361; Sargent C.A., Anderson M.J., Heisler S.L., Kendall E.,  
RA Gomez-Baccobar N., Campbell R.D.;  
RT "Characterization of the novel gene G11 lying adjacent to the  
RT complement C4A gene in the human major histocompatibility complex.";  
RL Hum. Mol. Genet. 3:481-488(1994).  
RN [11]  
RP SEQUENCE OF 1405-1431. AND SULTFATION.  
RX MEDLINE=86111851; PubMed=3944109; Horton G., Sims H., Strauss A.W.;  
RT "Identification of the site of sulfation of the fourth component of  
RT human complement.";  
RL J. Biol. Chem. 261:1786-1793(1986).

RN [12]  
RP STRUCTURAL BASIS OF POLYMORPHISM.  
RX MEDLINE=87080272; PubMed=2431902; Yu C.Y., Belt K.T., Giles C.M., Campbell R.D., Porter R.R.;  
RT "Structural basis of the polymorphism of human complement components  
RT C4A and C4B: gene size, reactivity and antigenicity.";  
RL EMBL J. 5:2873-2881(1986).  
RN [13]  
RP VARIANT C4A6 ALLOTYPES.  
RX MEDLINE=92242905; PubMed=1573268; Anderson M.J., Milner C.M., Cotton G.H., Campbell R.D.;  
RT "The coding sequence of the hemolytically inactive C4A6 allele of  
RT human complement component C4 reveals that a single arginine to  
RT tryptophan substitution at beta-chain residue 458 is the likely cause  
RT of the defect.";  
RL J. Immunol. 148:2795-2802(1992).  
RN [1]  
RP FUNCTION: C4 PLAYS A CENTRAL ROLE IN THE ACTIVATION OF THE  
CC CLASSICAL PATHWAY OF THE COMPLEMENT SYSTEM. IT IS PROCESSED BY  
CC ACTIVATED C1 WHICH REMOVE FROM THE ALPHA CHAIN THE C4A  
CC ANAPHYLATOXIN.  
CC [1]  
CC FUNCTION: DERIVED FROM PROTEOLYTIC DEGRADATION OF COMPLEMENT C4,  
CC C4A ANAPHYLATOXIN IS A MEDIATOR OF LOCAL INFLAMMATORY PROCESS. IT  
CC INDUCES THE CONTRACTION OF SMOOTH MUSCLE, INCREASES VASCULAR  
CC PERMEABILITY AND CAUSES HISTAMINE RELEASE FROM MAST CELLS AND  
CC BASOPHILIC LEUKOCYTES.  
CC [1]  
CC SUBUNIT: THIS PROTEIN IS SYNTHESIZED AS A SINGLE-CHAIN PRECURSOR  
CC AND, PRIOR TO SECRETION, IS ENZYMATICALLY CLEAVED TO FORM A TRIMER  
CC OF NONIDENTICAL CHAINS (ALPHA, BETA, AND GAMMA).  
CC [1]  
CC POLYMORPHISM: HUMAN COMPLEMENT COMPONENT C4 IS POLYMORPHIC WITH AT  
CC LEAST TWO LOCI, C4A AND C4B. 13 ALLELES OF C4A AND 22 ALLELES OF  
CC C4B HAVE BEEN DETECTED. THE ALLELE SHOWN HERE IS C4A4.  
CC [1]  
CC POLYMORPHISM: THE C4A ALLELES CARRY THE BLOOD GROUP RODGERS WHILE  
CC THE C4B ALLELES CARRY THE BLOOD GROUP CHIDO.  
CC [1]  
CC DISEASE: Defects in C4A are the cause of C4A deficiency  
CC [MIM:120810].  
CC [1]  
CC DISEASE: The C4A6 allele is totally deficient in hemolytic  
CC activity.  
CC [1]  
CC MISCELLANEOUS: C4A ALLOTYPES REACT MORE RAPIDLY WITH THE AMINO  
CC GROUP OF PEPTIDE ANTIGENS WHILE C4B ALLOTYPES REACT MORE RAPIDLY  
CC WITH THE HYDROXYL GROUP OF CARBOHYDRATE ANTIGENS.  
CC [1]  
CC MISCELLANEOUS: C4 IS A MAJOR HISTOCOMPATIBILITY COMPLEX CLASS-III  
CC PROTEIN.  
CC [1]  
CC SIMILARITY: TO C3, C5 AND ALPHA-2-MACROGLOBULIN.  
CC [1]  
CC SIMILARITY: Contains 1 anaphylatoxin-like domain.  
CC [1]  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: K02403; AAB59537.1; -  
CC EMBL: K02404; AAA59651.1; -  
CC EMBL: M59815; AAA51855.1; -  
CC EMBL: M59816; AAA51855.1; JOINED.  
CC EMBL: U24578; AAA9717.1; -  
CC EMBL: AF019413; AAB67980.1; -  
CC EMBL: AL049547; CAB89302.1; -  
CC EMBL: M14823; AAA55617.1; -  
CC EMBL: M14824; AAA52292.1; -  
CC EMBL: X77491; CAA54627.1; -  
CC PIR: I56095; C4HU.  
CC HSSP: P01031; 1KJ5.  
CC SWISS-2DPAGE: P01028; HUMAN.  
CC Genew: HGNC:1323; C4A.  
CC Genew: HGNC:1324; C4B.  
CC MIM: 120810; -  
CC MIM: 120820; -  
CC MIM: 120820; -  
CC CO: GO:000576; C:extracellular; NAS.  
CC GO: GO:0006958; P:complement activation, classical pathway; NAS.

DR GO: GO:0006954; P:inflammatory response; NAS.  
 DR GO: GO:0006937; P:regulation of muscle contraction; NAS.  
 DR InterPro: IPR002890; A2M\_N.  
 DR InterPro: IPR000020; Anaphylatoxin.  
 DR InterPro: IPR001599; Macroglobulin.  
 DR InterPro: IPR01134; Netrin\_C.  
 DR Pfam: PF00207; A2M; 1.  
 DR Pfam: PF01835; A2M\_N; 1.  
 DR Pfam: PF01821; ANATO; 1.  
 DR Pfam: PF01759; NTR; 1.  
 DR ProDom: PD003264; Anaphylatoxin; 1.  
 DR SMART: SM00104; ANATO; 1.  
 DR SMART: SM00643; C345C; 1.  
 DR PROSITE: PS00477; ALPHA\_2\_MACROGLOBULIN; 1.  
 DR PROSITE: PS01177; ANAPHYLATOXIN\_1; 1.  
 DR PROSITE: PS01176; ANAPHYLATOXIN\_2; 1.  
 KW Complement pathway; Plasma; Glycoprotein; Sulfation; Signal;  
 KW Inflammatory response; Polymorphism; Disease mutation;  
 KW Blood group antigen; Thioester bond.  
 FT SIGNAL 1 19  
 FT CHAIN 20 675 COMPLEMENT C4, BETA CHAIN.  
 FT PROPEP 676 679  
 FT CHAIN 680 1446 COMPLEMENT C4, ALPHA CHAIN.  
 FT PROPEP 1447 1453  
 FT CHAIN 1454 1744 COMPLEMENT C4, GAMMA CHAIN.  
 FT PEPTIDE 680 756 C4A ANAPHYLATOXIN.  
 FT DOMAIN 702 736 ANAPHYLATOXIN-LIKE.  
 FT DISULFID 703 728 BY SIMILARITY.  
 FT DISULFID 716 735 BY SIMILARITY.  
 FT CROSSLNK 1010 1013 Isoleucyl cysteine thioester (Cys-Gln).  
 FT MOD\_RES 1417 1417 SULFATION.  
 FT MOD\_RES 1420 1420 SULFATION.  
 FT MOD\_RES 1422 1422 SULFATION.  
 FT CARBOHYD 226 226 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 862 862 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1328 1328 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1391 1391 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT VARIANT 477 477 R -> W (in allotype C4A6).

Query Match 13.5%; Score 991.5; DB 1; Length 1744;  
 Best Local Similarity 24.7%; Pred. No. 2.8e-46;  
 Matches 409; Conservative 250; Mismatches 636; Indels 359; Gaps 65;  
 QY 26 PRFLVTPAGIRPGANVTIGVLELEHCHSOYTVAKELKTAS--NLTVSLVAGVPEKG 83  
 DB 21 PRLLFBSBVHLVGLPISVGVLODVRGQVYKGSVFLRNPSSKNVPCSPKVDTLSSER 80  
 QY 84 SFKTLTLPSPLPNSA-----DEIVELRYTGRTODELFPNST--RLSPETK 127  
 DB 81 DFALSL-QVPLKDAKSCGHLQLRGEVQLVAHSPMLKDSL--SRTNIGQINLFLSSR 137  
 QY 128 RISVFIQTDKALYKPKQEVKFRIVTLPSDFPKYTSNLILIKPKSNLIQOMLSQOSDLG 187  
 DB 138 RGHFLQTDQPIYVNGQRYRVRVFDQKMRPSTDTITVWENSHGLRV-----RKKEVY 192  
 QY 188 VIKSTFOLS-----SHPIIDGMSIQOVNDQTYGOS--FOVEEYVLKPEEYTLQRP-- 236  
 DB 193 MPSSIFODDEVIPDISP--GTWKISAFSDGLESNSTOFVEKVKYVLPENFEVKI--PFGK 249  
 QY 237 LYCGMNSKHLNG--TIRAKTYGPKVADVTLLPLSPFGKKKNI---TKTFKINGS 289  
 DB 250 PYILTVPHQHDMDLQARIYTKGPKGVAVYRFGLLDEGCKTFPFGLESQIKLVNGQ 309  
 QY 290 ANFSFNDKEMKNVDSN-GLSEYLDLSSPGPVELITVTESVTGISNVS--TNVFKQH 347  
 DB 310 SHISLSKAEFQDALEKIMGITDLOGLR---LVYAAIIRSPGEMEAELTSMYFVS 365  
 QY 348 DYILEFDTYTVAKPSLNFATVATRADGNQULIEERNNVTVYQRTNTEVSSGNS 407  
 DB 366 PFSJDLKTKRHLVGAFFLLQALVRENSGP-----ASGIPIKV-----SATVS 410  
 QY 408 GNGKMAVQKINTVPGSTFKIEFPILEDSESLQKAVFLGSKSMVAHSLFKSPS--K 465

DB 411 SPGSVPEVDIQONTDSSGGVSIPIIIPQTSISLQSVSAGSBHPALARLTVAAPSGCP 470  
 QY 466 TYIQKTRDEN-IKVGSPFL---VYSGNRKLELSYVVSROQLVAVG---KONSTWFS 518  
 DB 471 GFLSIKRPDSRPFRVGDITLNLINRAVSGATSEHYIIMLSRQIVPMNRPKRLTISVS 530  
 QY 519 LTPENSWTPRACVIVYIIEDDGEIISDLKIPVQL-VFKNKIKLYMSKVKA-EPESEVSL 576  
 DB 531 VFVDHILAPSFYVAFYHYGD-HPVANSILVDVQAGCEGKLEISVDGAKQYRNGESVKL 589  
 QY 577 RISVTPDLSIVGIVADKSVNLMAASN--DITMENVHELELVNTGYLLGMPNNSFVQ 634  
 DB 590 HLE-TJSLALVALGALDITALYAAASKHKPLNKGKFEANNSYDLGGPGGSAIQVFQ 648  
 QY 635 ECGP-----WYLT-----DNLTK-----DYI-----DGV- 654  
 DB 649 AAGLAFSDGQWTLSSRKRLSCPKEKTKRKKNVFOKALINEKIQYASPTAKCCQDGV 708  
 QY 655 -----YDAEYAEERFMEEN-----EGHIVD 674  
 DB 709 RLPMNRSCQRAARVQPDRCBPFLSCQFASLSLRKSRDKQAGLQRALEILOEDLID 768  
 QY 675 IHPFLSSPHVKHPETWIMDTMKYRI-----YQEPVVPDSITSWATGVISE 729  
 DB 769 EDDIP-----VRSFPEENLM-----RVEIVDRFQIITLWLPDLTWEIHGSLSK 815  
 QY 730 DLGIGLTTVEVLOAPRPFPIPLNPYSVRGEPFLIEITFNYLDATEVVKYIEKSK 789  
 DB 816 TKGL-CVATVQVLRFEFHLRLPMSVARFQLEBRPVLYNYLDKNTLVSHVSVVEG 874  
 QY 790 FDLIMTSSEINAGHOOTLVPSDEGATVLPRIPTHGEPITTVTALSP---TASDAVT 846  
 DB 875 LCL-----AGGGGLAQVLPVAGSARPAVSVPTAAAVSLKYVARGSEFPVGDAYS 928  
 QY 847 QMILVKAEGIEKYSQSIIDLTDNRLOSTLKLSE--SEPPYTVTGE---RVQITAI 901  
 DB 929 KVLQIEKEG--AIHRELVVEL--NPLDHRGRLTEIPGNSDPMMIIPDGDSNYVRVYASD 984  
 QY 902 --DYLRP-----SINGLASLRMPYGGGEOMINPAPVYIYLDLTKKKO---LTDNLKEX 952  
 DB 985 PLDTLSEGLALSPQVASTLRPRGCGEOTWITLAPLAASRYLDKTEQNSTLPPEYKDH 1044  
 QY 953 ALSEMRQYORELBYQGEDSPAFGNYDPGSGTWLSAPVLRGFLFLEADPYIDQVNLHR 1012  
 DB 1045 AVDLQKGVRIQGFRAQDSYAAWLSRD--SSTWLTAFVLKYLISLAQEVGSGPEKLOE 1102  
 QY 1013 TYTWLKGHQKSGEPMDPGKVIHSELQG--GNKSPVTLTAYIVTSL-IGRYKQP--- 1064  
 DB 1103 TSNWILSQOQADGSPQDPCFVLDRSMQGGVGNDEVALTAPVTIALHGLAVFQDEGAE 1162  
 QY 1065 -----NIDVQESIHPIESEFSRGISDNVTIALITYALSVSQP--KAKALMMLTWRAE 1116  
 DB 1163 PLKQRYEASTISKNSFEGEASAGLGAHAAATTAALSTKAPVLDLGVAAHNNLMAAQ 1222  
 QY 1117 QEGGMQFWSSEKLSDSMQPR-----SLDEVAAYALSHFLOFQTSSEGI 1162  
 DB 1223 ETGDNLYMSVYTSQSNASPTAPRNPDPMPQAPALMLETAYALL-HLL---LHEGK 1278  
 QY 1163 PIM-----RWLSQRNSLGGFASTODTTVALKALSEF-AALNMTERTNIOVTVGSPSS 1216  
 DB 1279 AEMADQASAWLTRQSGFQGGFRSTOPTVALDALSAWIMASHHTTEEGCLAVTIS- 1332  
 QY 1217 PLAVVQPMANVISANGRAICQLN-----VYVNYKASSSSRRRSI-- 1258  
 DB 1333 -----STGRNGKSHALQNNRQJRLGLEELQFSLGKINNVGVGNSGKTLKVL 1382  
 QY 1259 -----QNGEAPDLVAV-----KENKDLNHDV- 1282  
 DB 1383 TYNVLDKNTKTCODLQIEVTVKGVHVEYTMENANDYEDYEDLPADDDPAPLPQVTPLO 1442  
 QY 1283 -----NVCTSFSGP--GRSGMALMEVNLISGFNV---PS 1311



DB 1443 LEBGRNRRRRAPKVEQESRWHYTCIMRNKGVLSCMAIADVTLLSGFHALRADIE 1502

QY 1312 EAISLSET-VKKVEYDHGKLTLYLDSVNETOPCVNIPAVRNFXVNTQDASVSIYDYYP 1370

DB 1503 KLTSLSDRYVSHPEHTEGPHVLLYFDSVPFTRSECVGEANQVEPVGLVQASATLYDYNYP 1562

QY 1371 RQAVRSVNSEYK-----LSSCDLCSVDQCCRP 1398

DB 1563 ERRCSVFYGAPSKSRLLATLCSAEVQCAEGKCP 1596

RESULT 11

CO4\_MOUSE STANDARD; PRT; 1738 AA.

AC P01029; Q61859; 21-JUL-1986 (Rel. 01, Created)

DT 01-AUG-1988 (Rel. 08, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Complement C4 precursor [Contains: C4A anaphylatoxin].

GN C4.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;

OC NCBI\_TaxID=10090;

RN (1)

RP SEQUENCE FROM N.A.

RX MEDLINE=85298264; PubMed=3862104;

RA Sepich D.S., Noonan D.J., Ogata R.T.;

RT "Complete cDNA sequence of the fourth component of murine

RT complement.";

RL Proc. Natl. Acad. Sci. U.S.A. 82:5895-5899 (1985).

RN (2)

RP SEQUENCE FROM N.A.

RC STRAIN=B12.WR;

RX MEDLINE=87309760; PubMed=3624865;

RA Rosa P.A., Sepich D.S., Robins D.M., Ogata R.T.;

RT "Constitutive expression of S1p genes in mouse strain B10.WR directed

RT by C4 regulatory sequences.";

RL J. Immunol. 139:1568-1577 (1987).

RN (3)

RP SEQUENCE FROM N.A.

RC STRAIN=B12.WR; TISSUE=Liver;

RX MEDLINE=89380278; PubMed=2777798;

RA Ogata R.T., Rosa P.A., Zepf N.E.;

RT "Sequence of the gene for murine complement component C4.";

RL J. Biol. Chem. 264:16565-16572 (1989).

RN (4)

RP SEQUENCE FROM N.A.

RC STRAIN=FM; TISSUE=Liver;

RX MEDLINE=85289294; PubMed=2993295;

RA Nonaka M., Nakayama K., Yeul Y.D., Takahashi M.;

RT "Complete nucleotide and derived amino acid sequences of the fourth

RT component of mouse complement (C4). Evolutionary aspects.";

RL J. Biol. Chem. 260:10936-10943 (1985).

RN (5)

RP SEQUENCE OF 651-810 AND 924-1083 FROM N.A.

RX MEDLINE=85038607; PubMed=6208559;

RA Nonaka M., Takahashi M., Natsume-Sakai S., Nonaka M., Tanaka S.,

RA Shimizu A., Honjo T.;

RT "Isolation of cDNA clones specifying the fourth component of mouse

RT complement and its isotype, sex-limited protein.";

RL Proc. Natl. Acad. Sci. U.S.A. 81:6822-6826 (1984).

RN (6)

RP SEQUENCE OF 1099-1142 FROM N.A.

RX STRAIN=B10.WR, C57BL/6, C3H/He, CBA/J, B10.BR, and DBA2;

RX MEDLINE=90533398; PubMed=2387317;

RA Ogata R.T., Zepf N.E.;

RT "C4 from C4-high and C4-low mouse strains have identical sequences in

RT the region corresponding to the isotype-specific segment of human

RT C4.";

RL Eur. J. Immunol. 20:1607-1610 (1990).

RN (7)

RP SEQUENCE OF 1105-1449 FROM N.A.

RX MEDLINE=85166208; PubMed=3856857;

RA Levi-Strauss M., Toei M., Steinmetz M., Klein J., Meo T.;

RT "Multiple duplications of complement C4 gene correlate with H-2-

RT controlled testosterone-independent expression of its sex-limited

RT isoform, C4-S1p.";

RL Proc. Natl. Acad. Sci. U.S.A. 82:1746-1750 (1985).

RN (8)

RP SEQUENCE OF 1257-1376 FROM N.A.

RX MEDLINE=85038859; PubMed=6149581;

RA Toei M., Levi-Strauss M., Duponchel C., Meo T.;

RT "Sequence heterogeneity of murine complementary DNA clones related to

RT the C4 and C4-S1p isoforms of the fourth complement component.";

RL Philos. Trans. R. Soc. Lond., B, Biol. Sci. 306:389-394 (1984).

RN (9)

RP SEQUENCE OF 1360-1511 FROM N.A.

RX MEDLINE=83273751; PubMed=6192448;

RA Ogata R.T., Shreffler D.C., Sepich D.S., Lilly S.P.;

RT "cDNA clone spanning the alpha-gamma subunit junction in the

RT precursor of the murine fourth complement component (C4).";

RL Proc. Natl. Acad. Sci. U.S.A. 80:5061-5065 (1983).

RN (10)

RP SEQUENCE OF 1-128 FROM N.A.

RC STRAIN=FM; TISSUE=Liver;

RX MEDLINE=86031969; PubMed=2997024;

RA Nonaka M., Nakayama K., Yeul Y.D., Shimizu A., Takahashi M.;

RT "Molecular cloning and characterization of complementary and genomic

RT DNA clones for mouse C4 and S1p.";

RL Immunol. Rev. 87:81-99 (1985).

RN (11)

RP SEQUENCE OF 1-21 FROM N.A.

RX MEDLINE=87017050; PubMed=3464002;

RA Nonaka M., Kimura H., Yeul Y.D., Yokoyama S., Nakayama K.,

RA Takahashi M.;

RT "Identification of the 5'-flanking regulatory region responsible for

RT the difference in transcriptional control between mouse complement C4

RT and S1p genes.";

RL Proc. Natl. Acad. Sci. U.S.A. 83:7883-7887 (1986).

CC -!- FUNCTION: C4 PLAYS A CENTRAL ROLE IN THE ACTIVATION OF THE

CC CLASSICAL PATHWAY OF THE COMPLEMENT SYSTEM. IT IS PROCESSED BY

CC ACTIVATED C1 WHICH REMOVE FROM THE ALPHA CHAIN THE C4A

CC ANAPHYLATOXIN.

CC -!- SUBUNIT: THIS PROTEIN IS SYNTHESIZED AS A SINGLE-CHAIN PRECURSOR

CC AND, PRIOR TO SECRETION, IS ENZYMATICALLY CLEAVED TO FORM A TRIMER

CC OF NONIDENTICAL CHAINS (ALPHA, BETA, AND GAMMA).

CC -!- MISCELLANEOUS: C4 IS A MAJOR HISTOCOMPATIBILITY COMPLEX CLASS-III

CC PROTEIN.

CC -!- SIMILARITY: TO C3, C5 AND ALPHA-2-MACROGLOBULIN.

CC -!- SIMILARITY: Contains 1 anaphylatoxin-like domain.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>

CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC -----

DR EMBL; K00019; AAA39554.1; -

DR EMBL; M11729; AAA39506.1; -

DR EMBL; M12968; AAA39558.1; -

DR EMBL; M12970; AAA39555.1; -

DR EMBL; M12972; AAA39556.1; -

DR EMBL; M12969; AAA39559.1; -

DR EMBL; M11789; AAA39557.1; -

DR EMBL; K02798; AAC42021.1; -

DR EMBL; M17440; AAA39561.1; -

DR EMBL; M14225; AAA39563.1; -

DR EMBL; M14226; AAA39684.1; -

DR EMBL; X55493; CAA39112.1; -

DR EMBL; X55495; CAA39114.1; -

DR PIR; A24558; A24558.

DR PIR; A29176; A29176.

DR HSPB; P01031; 1KJ5.

DR MGD; MGI:88228; C4.  
 DR InterPro: IPR002890; A2M\_N.  
 DR InterPro: IPR000020; Anaphylatoxin.  
 DR InterPro: IPR001599; Macroglobulin2.  
 DR InterPro: IPR001134; Netrin\_C.  
 DR Pfam: PF00207; A2M; 1.  
 DR Pfam: PF01835; A2M\_N; 1.  
 DR Pfam: PF01821; ANATO; 1.  
 DR Pfam: PF01759; NTR; 1.  
 DR ProDom: PD003264; Anaphylatoxin; 1.  
 DR SMART: SM00643; ANATO; 1.  
 DR SMART: PS00477; ALPHA\_2\_MACROGLOBULIN; 1.  
 DR PROSITE: PS01177; ANAPHYLATOXIN\_1; 1.  
 DR PROSITE: PS01178; ANAPHYLATOXIN\_2; 1.  
 KW Complement pathway; Plasma; Glycoprotein; Sulfation; Signal;  
 KW Inflammatory response; Thioester bond.  
 FT SIGNAL 1 19  
 FT CHAIN 20 673 COMPLEMENT C4, BETA CHAIN.  
 FT PROPEP 674 677 COMPLEMENT C4, ALPHA CHAIN.  
 FT CHAIN 678 1443  
 FT PROPEP 1444 1447  
 FT CHAIN 1448 1738  
 FT PEPTIDE 678 753 COMPLEMENT C4, GAMMA CHAIN.  
 FT DOMAIN 700 734 ANAPHYLATOXIN-LIKE.  
 FT DISULFID 700 726 BY SIMILARITY.  
 FT DISULFID 701 733 BY SIMILARITY.  
 FT DISULFID 714 734 Iso-glutamyl cysteine thioester (Cys-Gln  
 FT CROSSLINK 1006 1009 (by similarity).  
 FT MOD\_RES 1413 1413 SULFATION.  
 FT MOD\_RES 1416 1416 SULFATION.  
 FT MOD\_RES 1417 1417 SULFATION.  
 FT CARBOHYD 224 224 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 743 743 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1387 1387 F->Y (IN REF. 4).  
 FT CONFLICT 132 132 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 327 327 G->Y (IN REF. 4).  
 FT CONFLICT 327 327 G->E (IN REF. 4).  
 FT CONFLICT 570 570 Q->E (IN REF. 4).  
 FT CONFLICT 720 720 R->G (IN REF. 5).  
 FT CONFLICT 739 740 DL->A (IN REF. 5).  
 FT CONFLICT 838 838 P->R (IN REF. 4).  
 FT CONFLICT 933 933 P->L (IN REF. 5).  
 FT CONFLICT 1043 1043 D->E (IN REF. 5).  
 FT CONFLICT 1119 1119 V->A (IN REF. 7).  
 FT CONFLICT 1190 1190 A->T (IN REF. 7).  
 FT CONFLICT 1324 1324 K->N (IN REF. 4).  
 FT CONFLICT 1401 1401 G->S (IN REF. 9).  
 FT CONFLICT 1442 1442 R->K (IN REF. 4).  
 FT CONFLICT 1453 1453 A->V (IN REF. 4).  
 SQ SEQUENCE 1738 AA; 192870 MW; D1E02AE7AB42BFF CRC64;  
 Query Match 13.4%; Score 987; DB 1; Length 1738;  
 Best Local Similarity 24.9%; Pred. No. 4.8e-46;  
 Matches 415; Conservative 245; Mismatches 643; Indels 366; Gaps 64;  
 QY 10 AHLICVCAALAVAPGPRFLVTAPGIRPGNVTIGVLEHCSQYTVKAE----- 62  
 DB 8 AMVVSFCASSL---QKPRLLFPSVSVNLGTPLSVGQQLDAPGQ-EVKSQVFLRMPKG 63  
 QY 63 -----LKTASNLTVSVLEA--EGVFESKFKTLTLPSPILNSADELYELRVGT 110  
 DB 64 GSCSPKDFKLSGGDFVLISLEVLVEDVRSGLFDLRBAPHILQVAGSP--WLRLNAPK 121  
 QY 111 QDELIFSNSTLSEETKRIISVFIOTDKALYKPKQEVKFRIVTLFSDFKPYKTSINILIKD 170  
 DB 122 ATE---TQGVNVLFPSSRGHIFVQTDQPIYNPGQVRVRYVPALQKKRPSPTDFLTIVEN 178  
 QY 171 PKS-NLIQOMLSQSDGIVISKTFQLSHPILGDMISIQOVND---QTYVQSPQVSEVYL 226  
 DB 179 SHGRVLLKKEIFETISIFQDAPFTIPDISEP--GTMKLSARPSDGLSENRSTHFEVKKTVL 236  
 QY 227 PKFEVTLQTP-----LYCSMNSKHLNGITITAKYTYGKPVKGDVLTLEFLPSFMGKKINI- 280

DB 237 PNFEVKI-TPWKPYILLVPSNSDEIQLDIQRATYIGKPVGVAVYTRPALMDEQKTFELR 295  
 QY 281 ---TKTEFKINGSANFSFNDENMGVMDSSN-----GLSEYLDLSPGFEVILTYTES 330  
 DB 296 GLETOAKLVGRGRHISISKQFOALDKINIGRDLLEGLRY-----AATAVIES 345  
 QY 331 VTGISRNVS--TNVFPKQHDYIEFFDYTVLKESLNFATVATVKTADGNQLTLEBRNNV 389  
 DB 346 PGSEMERAEELTSWRFPSSAFSLDSRTKRLVPGAHFLQALVQEMSGSE-----ASN 399  
 QY 390 VITVYQNTYENYSGNSGNGKQKAEAVQKINYTVTPOSGTFIEEPILEDSELOKAYFLG 449  
 DB 400 PVKVS-----ATLVSGDS--QVLDIQSTV---GIGQVSISEPIPTVTELRLLVSAGS 449  
 QY 450 SKSMVAHSLFKSPSK--TYIOLKTRD-ENIKVSGFPEL-----VSGNKLKELSYVVS 502  
 DB 450 LYPALRLTVQAPPSKGTGLSIEPDRPSVGDFTILNLQVGVGAPAPFESHYYMIIIS 509  
 QY 503 RGQLVAVGKO---NSTWFSLTTPNSMTPKACVIVYIIEDDGEIISDVLKIPVO---LVFK 556  
 DB 510 RGQIMMAGREPRKTVTSVSLVDHQLASFYFAVYFH--QGHVANSLLINISRDCEGK 568  
 QY 557 NKIKLWYSKYKAPSEKVSRLISVTOQDSIVGIVAVDKSVNLM--NANSNITMENYVHEL 614  
 DB 569 LQKLVGAK-EYRNADMVKRLIQ-TDSKALVALGAVDTALVAVGSRSHKLDMSKVFEVI 626  
 QY 615 ELVNTGYLLFMNPSFAVPOEGLMVLTDLANTKDYID----- 652  
 DB 627 NSTNVGCGPGGDALQVFDAGLAFSDGDRLTQTRBEDLSCPEKKSQRQKRVNFOKAVS 666  
 QY 653 ---GVYDN----- 667  
 DB 687 EKLGGVSSPPARKCCQDMKTLPMKRTCEQRAARVPOACREPFISCKKABEDLRNQT 746  
 QY 668 NEGH-----YVDIHDPSSGSEPHVKHPEWTW---LDTNMGRIYIOEFE 710  
 DB 747 SQAHLRANNHMLQEBDLIDEDDIL-----VRTSPENMLMVEPVDSS-----KLLT 794  
 QY 711 VTPDSTSVVATGPIYSIEDLGILTTTPVELQAFQFFIFLNLPSVIRGSEFPALEITI 770  
 DB 795 VMLPDSMTTWELHGVSSKSKGL-CVAKPTRVAVFRFHILHLPLSIRRFQEBELRPVL 853  
 QY 771 FNYLKATEVYKVIKESDKEDILMTSEINATGHQOQLLVPSDEGATVLPPIRPHLGEI 830  
 DB 854 YNVLND--DVAVSVHTPVSGGLAGGMA-----QGVTPASARVAVASVPTAAV 907  
 QY 831 PIVVTALSP-TASDAYTQMLIVKAGIEKSYQSILDLTD-NRLOSTLKLTSFSPFN- 887  
 DB 908 PLRVVARGVFDLGDVASKILOIEKEG--AIHRELVYINLDPILNLLGRITLE-IPGSSDPNI 964  
 QY 888 -----TYMGSERVQITAGDVLGPIINGLASLIRMPVCGCGGOMNINFAPIYILD 937  
 DB 965 VPDGPRSSLVRYVASPELMTGSEBALSP--GGVASLRLRPQCCAQOTMYLAPTLASN 1022  
 QY 938 YLTKKKO---LITNLLEKALSPKOCYORELLAQREDSGSARANYDPGSGTWLSAFVLR 994  
 DB 1023 YLDRTEGMSKLSFETDHDVNDILOKGYMRLQGRKKNDSGAMLRHD--STWLTAFVLK 1080  
 QY 995 CLEADPYIDIDONVLRHTYTWLKGHOKSNGEFPWDPGRVHISLQG---GNKSPVLTAV 1051  
 DB 1081 ILSLAQOVNSPEKLOETASWLLAQQLGDSFHDPCPVLRHAMQGLVGSDETVALTAF 1140  
 QY 1052 IYVSL-----LGRKYQPNIDVQESIHLELSESPRGISDNVYTLALITYALSSVG 1100  
 DB 1141 VVALHGHGIDVODDDAKQKNNVEASITVANSFLQOKASAGLLGAHAAAITVALTLT- 1199  
 QY 1101 SPKAKAL-----NMLTWRADEGGMQFW---VSSESK-LSDSWQPRS-----LD 1141  
 DB 1200 --KASEDLRVAVANSMAAAEFEGHILYNGLVGSGQDKVLRRTARRSPPEVPQAPALW 1257  
 QY 1142 IEVAAYALLSHFLQFQTSSEGPIIM-----RWLSGRQNSLGGFASSTQDTYVALKALSEF-A 1195

DR	PfAm; PFO1835; A2M.N; 1.
DR	PfAm; PFO1821; ANATO; 1.
DR	PfAm; PFO1759; NTR; 1.
DR	PrDom; PD003264; Anaphylatoxin; 1.
DR	SMART; SMO0104; ANATO; 1.
DR	SMART; SMO0643; C3A5C; 1.
DR	PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
DR	PROSITE; PS01177; ANAPHYLATOXIN 1; 1.
DR	PROSITE; PS01178; ANAPHYLATOXIN 2; 1.
KW	Complement pathway; Complement alternate pathway; Plasma;
KW	Inflammatory response; Glycoprotein; Thioester bond.
FT	NON_TER 1 1
FT	CHAIN 1 1640 COMPLEMENT C3.
FT	CHAIN 1 642 BETA CHAIN.
FT	CHAIN 647 1640 ALPHA CHAIN.
FT	PEPTIDE 647 722 C3A ANAPHYLATOXIN.
FT	CHAIN 723 1640 C3B (ALPHA CHAIN).
FT	PEPTIDE 723 931 C3C FRAGMENT (BY SIMILARITY).
FT	PEPTIDE 932 1278 C3D FRAGMENT (BY SIMILARITY).
FT	PEPTIDE 932 1033 C3E FRAGMENT (BY SIMILARITY).
FT	PEPTIDE 1034 1278 C3F FRAGMENT (BY SIMILARITY).
FT	PEPTIDE 1279 1295 C3F FRAGMENT (BY SIMILARITY).
FT	SITE 722 723 CLEAVAGE (BY C3 CONVERTASE).
FT	SITE 931 932 CLEAVAGE (BY FACTOR I) (BY SIMILARITY).
FT	SITE 1278 1279 CLEAVAGE (BY FACTOR I) (BY SIMILARITY).
FT	SITE 1295 1296 CLEAVAGE (BY FACTOR I) (BY SIMILARITY).
FT	DOMAIN 668 703 ANAPHYLATOXIN-LIKE.
FT	DOMAIN 1420 1430 PROPERDIN-BINDING.
FT	DISUFLD 536 797 INTERCHAIN (BY SIMILARITY).
FT	DISUFLD 603 638 BY SIMILARITY.
FT	DISUFLD 668 692 BY SIMILARITY.
FT	DISUFLD 669 702 BY SIMILARITY.
FT	DISUFLD 682 703 BY SIMILARITY.
FT	DISUFLD 853 1488 BY SIMILARITY.
FT	DISUFLD 1079 1135 BY SIMILARITY.
FT	DISUFLD 1335 1464 BY SIMILARITY.
FT	DISUFLD 1481 1486 BY SIMILARITY.
FT	DISUFLD 1493 1563 BY SIMILARITY.
FT	DISUFLD 1510 1638 BY SIMILARITY.
FT	DISUFLD 1614 1623 BY SIMILARITY.
FT	CARBOHYD 164 164 N-LINKED (GLCNAC..).
FT	CROSSLNK 988 991 IsoglutamyI cysteine thioester (Cys-Gln)
SO	SEQUENCE 1640 AA; 182104 MW; 0965164PAFLB87812 CRC64;
Query Match	13.4%; Score 964.5; DB 1; Length 1640;
Best Local Similarity	24.1%; Pred. No. 6.1e-46;
Matches 387; Conservative 298; Mismatches 571; Indels 349; Gaps 67	
Oy	18 ALALAAGPRPLVLTAGIIRPGAGNTIGIVELLEHCPSGVYKRLTATSNLTVSVLEAE 77
Dd	1 AALQV-----LSANLLRGSNNENITVESODHGGELANKL-MVKNHHPQSKELSKS 52
Oy	78 GVFEK-GSEFKTLT-----LPSLPLNSADEIELRYLRGTODEILFSNSTRLSFEETR 128
Dd	53 VVLDDQANNQANTOLVIQRGPLVDDPKRKQYVVLQAQEPRLLEKVL-----VSPOSGV 107
Oy	129 ISVFIOETKALYKPQOEKFRIVLTESDFKY-----KTSNLILIKDKPSN 174
Dd	108 I--FIOTDKTIYTPASTVHYRVFSMTPLGEPLREIFEDEQVANKEIAVSVIEIMPENI 165
Oy	175 LIQQMLSQSDLDGVIKTFQLSSHHITLDMGIQOVN---DQTYYSOFVSEVYLPPFEV 231
Dd	166 TIFREL-VNPDDGVAISGGQKLPIDIVSFTHMVIVRRPOSTPQKTSPSEFEVEKVELPSFEV 224
Oy	232 TLQTP--LYCSNMWSKHNGITAKTYTKPKRGCVTLTF-----LPLSFMG--- 275
Dd	225 SL-TPKAKAFYYDDNDLVIDITARLYKGKVTGIGYVFGVITTESKKSPASLQKVEI 283
Oy	276 -----KKKNITKTP-KINGSANFSFNDEMKNVWDSSNGISEVYLDSPPQVEILT 325
Dd	284 KQGKVACLKKEHIQTQFPKIHDLVKOSI-----FVSVSVLTEGGGEVNAEKRG-IQIVT 338

Qy	326	TVTESVTGISRNVTNPFKQHDYIIIEFDYTYTLKPSLNFATKATKVRADQNGQLTEBR	385
Db	339	S-----PYLFRKTRPKYFKPCGMPVUSYITINP-----	368
Qy	386	RNNVITVTOBNYTEWGSNSGNOKMEAVOKINYTPQSGTFKIEPFLBEDSELQKA	445
Db	369	-NSPAIGVEVETPDHAKGVTRAN-----GPAKIPLMTVASATELVITV	411
Qy	446	-----YFLGSSSMAVHSLFSPSKTYIQKTRDENIVGSPPELVUS-GNKKRL--	493
Db	412	KTKPDGDRQOTGGGTMRALP--YRTSTKNFLHVDSDNELGIDPIKIDMLGPTTIPN	469
Qy	494	KELSYNVSRGOLVAVG---KONSTMFSLT--PENSMTPKACVIVYIYLEDGEIISDYLK	548
Db	470	HDLTMTFLSRGOLVAVGFRKQGNALVTLISVPSKELLPSFRTIYVHYGADVLADSVW	529
Qy	549	IPVOLFRKNKIKLYMSKVKA--EPSEKYSLRISVTPQPSIVIGIVAVDKSVMLMANSNDIT	606
Db	530	VDIKVSCMGSLKVTSTRPASVEPRARSLTIT--GDPGAKVGLVAVDKGVVYLANSKHRLT	588
Qy	607	MENVVHELELYTGYLGMFANSPAVFOECGL-----WVLTDA--	645
Db	589	QTKIWDLIEKHDTGCTAGGADNMGVFYDAGLVPELTNAKTGRTIDPSCVPSRRRAV	648
Qy	646	-----LTKD-YIDGYDN-----AERAE	662
Db	649	TISDYITSMASKYHGLAKECCVDGRDNTMGYTCRRRAQYISDGVCVQALFVCTEMAS	708
Qy	663	RFMEBNEGHIV-----DIHDFSLGSSPHV-RKHPETWILDTNMGVRIYQ--EFEVT	712
Db	709	KKIEKQOALLNRSREDDDDAVYRSBDIYRSQFPESMMEDINLECPAQNHGHCST	768
Qy	713	-----VPDSITSVATGTVISEDLGLGTLTPVELQAFQDFITLMLPISVINGEPA	765
Db	769	SVIRNNFLKDSITTQITAITAISLKHGFI-CVADPEMIVLKEFFIDLKLPYSAVNEOLE	827
Qy	766	LEITFNTLKDATERKVIIEKSDKFDILMTSEINATHQOILLVPSDEATVLEPIRPT	825
Db	828	VKAILAHNSEDPTIYKVELMENG-----VCSASKKGYKROCVANDPMSTRVVPYIIPM	883
Qy	826	HLG-EIITVTVALSPASDAVTOMLVKAEGIEKYSQSILDLTDRLQSTLTKLFS	883
Db	884	KLGLHSIEVKA SVKXSGSDGVKRLRVABEVLVKKGTNVLLNVKHKGEQT-SHISG	942
Qy	884	FPENTVTGSR--VOITAIQD---VLGPSING--LASLIRMPYGCBBQNMINFANITY	934
Db	943	VPRNQVPSNDAADTLISVTA-GEQTSVLVEQAIISGDSLIAVQVPGCEGQMMIYMTLEVI	1001
Qy	935	ILDYITKKQOLD---NLKEXALSTMROGYQRELLYQREDSFSAFGNYDPBGSTWLSAF	991
Db	1002	ATHYIDNTKKMEDIGLDKRNTRAIKINIGYQOLAVRKREDSYAAW--VSRQSTWTLRAY	1055
Qy	992	VLRCFLIADBPYIDIDONYLHRTYTWL-KQHQRKSEFMDPGRHISELQG--GNKSPVT	1047
Db	1060	VVKVPAMSTLISQVENULCTRAVKLLINTQOPDIEFHEFAVHIAETGMVNRGSDNDAS	1119
Qy	1048	LTAYI-----VTSLLGKRYKYPNIDVQESIHFLSEFSRGISDNYTLALIT	1093
Db	1120	MTAFVLLIMQEASSVCEQSVNSLPG-----SMAKAVAYLEKRLPH-LTNPVAVAMTS	1170
Qy	1094	YALSSVSGPKAKKALNMLTMRAEQ--EGGQMFVWSSSKLSDSQPSLDIEVAAYALL	1155
Db	1171	YALNAGLKNKETLTKFASPOLDHMPVDBGYQY-----TLEATSYALL	1213
Qy	1151	SHPLQFOTSEGIPIRWLTSRONSLJGFSASTODTVALKALSEFALNMNTER--TNIQ	1207
Db	1214	ALVYKPAEEKGPIVKMLNKQKKQGGGSGISTOSTIMVFQVAEYISHYKDKDFPLNINVL	1273
Qy	1208	TVTGPSSBSPILAVVQPMANVISANGFGPAICOLNV---YNYKASGSSRRRRSI-----	1258
Db	1274	EVAGRAS-----VTKMSIN-NKQGFPHIRTRKVNISIDDLQVTKKASGNEATLISVTLVYA	1326
Qy	1259	-----QNGEAPDLDVAVKENKDINHVD-----LNVCTSESGFGRSQ-MALNEVNILS	1305

```

Db      1327 LPEKSDSCBESFDLSVLT-K-MDRTSHEDAKESFMILITVULYKNSSEDAITMSIDILLT 1365
Qy      1306 GFMWPSERIS-LSE-----TYKKVEYD-----HGKLNLYDSVN--ETQFCVNIPAVRN 1352
Db      1386 GFIVDTDLNLQTSKGRERYIEKFEPMDEKVLSEKSGSLIYLDKVSHEKDEDRISFKIHRVO-- 1443
Qy      1353 KVSNTQDASVSIVDYEPFRQAVNSYSE-----VLTSSDCIS 1391
Db      1444 EVGVLPAAVSVYEYNOKR-CVKFYHPQREBGTLISRLCIADVCT 1487

RESULT 13
CO3_NAJNA
ID CO3_NAJNA STANDARD; PRT; 1651_AA.
AC Q01833;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Complement C3 precursor [contains: C3a anaphylatoxin].
C3.
OS Naja naja (Indian cobra).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidoptera; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Elapidae; Elapinae; Naja.
NCBI_TaxID=35670;
(1)
SEQUENCE FROM N.A.
RP TISSUE=Liver;
RX MEDLINE=93056528; PubMed=1431125;
RX Fritzlinger D.C., Connolly M., Petrella E.C., Bredenhorst R.,
RA Vogel C.W.;
"Primary structure of cobra complement component C3.";
RT J. Immunol. 149:3554-3562(1992).
RL
CC -1- FUNCTION: C3 PLAYS A CENTRAL ROLE IN THE ACTIVATION OF THE
CC COMPLEMENT SYSTEM. ITS PROCESSING BY C3 CONVERTASE IS THE CENTRAL
CC REACTION IN BOTH CLASSICAL AND ALTERNATIVE COMPLEMENT PATHWAYS.
CC AFTER ACTIVATION C3B CAN BIND COVALENTLY, VIA ITS REACTIVE
CC THIOLESTER, TO CELL SURFACE CARBOHYDRATES OR IMMUNE AGGREGATES.
CC -1- FUNCTION: DERIVED FROM PROTEOLYTIC DEGRADATION OF COMPLEMENT C3,
CC C3A ANAPHYLATOXIN IS A MEDIATOR OF LOCAL INFLAMMATORY PROCESS. IT
CC INDUCES THE CONTRACTION OF SMOOTH MUSCLE, INCREASES VASCULAR
CC PERMEABILITY AND CAUSES HISTAMINE RELEASE FROM MAST CELLS AND
CC BASOPHILIC LEUKOCYTES.
CC -1- SUBUNIT: C3 precursor is first processed by the removal of 4 Arg
CC residues, forming two chains, beta and alpha, linked by a
CC disulfide bond. C3 convertase activates C3 by cleaving the alpha
CC chain, releasing C3a anaphylatoxin and generating C3b (beta chain
CC + alpha' chain) (By similarity).
CC -1- SIMILARITY: TO C4, C5 AND ALPHA2-MACROGLOBULIN.
CC -1- SIMILARITY: Contains 1 anaphylatoxin-like domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; L02365; AAA49385.1; -.
DR HSSP; P01024; IC3D.
DR InterPro; IPR002890; A2M_N.
DR InterPro; IPR000020; Anaphylatoxin.
DR InterPro; IPR001840; Anaphylatoxin.
DR InterPro; IPR001599; MacroglobulinA2.
DR InterPro; IPR001134; Netrin_C.
DR Pfam; PF00207; A2M; 1.
DR Pfam; PF01835; A2M_N; 1.
DR Pfam; PF01821; ANATO; 1.
DR Pfam; PF01759; NTR; 1.
DR PRINTS; PRO00004; ANAPHYLATOXN.
DR PRODom; PD003264; Anaphylatoxin; 1.

```

DR	SMART; SMO0104; ANATO; 1.	
DR	SMART; SMO0643; C345C; 1.	
DR	PROSITE; PS00477; ALPHA 2. MACROGLOBULIN; 1.	
DR	PROSITE; PS01177; ANAPHYLATOXIN 1; 1.	
DR	PROSITE; PS01178; ANAPHYLATOXIN 2; 1.	
KW	Complement pathway; Complement alternative pathway; Plasma; Inflammatory response; Signal; Thioester bond.	
FT	SIGNAL	1 22
FT	CHAIN	23 1651
FT	CHAIN	23 655
FT	CHAIN	661 1651
FT	PEPTIDE	661 738
FT	CHAIN	739 1651
FT	SITE	738 739
FT	DOMAIN	683 718
FT	DISULFID	546 807
FT	DISULFID	615 650
FT	DISULFID	683 710
FT	DISULFID	684 717
FT	DISULFID	697 718
FT	DISULFID	863 1501
FT	DISULFID	1091 1147
FT	DISULFID	1346 1477
FT	DISULFID	1377 1446
FT	DISULFID	1494 1499
FT	DISULFID	1506 1578
FT	DISULFID	1525 1649
FT	CROSSLINK	999 1002
SEQUENCE	1651 AA; 184926 MW; 66691BD8F4FA935F7 CRC64; (by similarity).	
Q	Isoglutamyl cysteine thioester (Cys-Gln).	
Q		

Query Match	13.2%;	Score 973.5;	DB 1;	Length 1651;
Best Local Similarity	23.5%;	Pred. No. 2.5e-45;		
Matches 391;	Conservative 278;	Mismatches 649;	Indels 343;	Gaps 61;

```

QY 1 MOGPBLTAHLVCYCTAALAVAPGPRFLVTPAGIIRPGANVTIGVELL-EHCPSQVTVK 59
  |:::|
  |:::|
  |:::|
Db 1 MEGMALYLVAALLIIFPGS---SHGALYLLITPRVLRDTBEQILVEAHGDSIPKSLDIF 57
  |:::|
  |:::|
  |:::|
QY 60 A----ELKLTASNLTYSYLEAGVEKGSFKLTLPLSLPLNSADELYELRVGRIDDEL 115
  |:::|
  |:::|
  |:::|
Db 58 VHDEPKOKTLEQSRHVNQAGSMFVT---PIIKYPAKELINDSKONQYVVKVGPQVA 114
  |:::|
  |:::|
  |:::|
QY 116 FSNSTRLEFETRIKISVFIOTDKALKPKQCEYFRJLYLSPDEKPYKTSJLNLILIKPKSWL 175
  |:::|
  |:::|
  |:::|
Db 115 LEKVLLYSQDS--FVFIOTDKGITYPGSPVRIRVFSV--DHNHMDKTVIVE----- 164
  |:::|
  |:::|
  |:::|
QY 176 IQOMLSQOSDLGIVISKTFOLSSHPI-----LGDWSIQOVN---DQTY 216
  |:::|
  |:::|
  |:::|
Db 165 -----FQTPREIV-----VSSKPPNPGSIRPNYLPVLVSCTYKAQVAKYHSPRESTT 213
  |:::|
  |:::|
  |:::|
QY 217 QSPQVSEYVLPKEVYTLQTP---LYCSMNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSF 273
  |:::|
  |:::|
  |:::|
Db 214 AYFDREYVLPBEFVRLQPSDKFLYIDGN-KAFHVSITARYLYGKVEG-VAFVPGVKI 271
  |:::|
  |:::|
  |:::|
QY 274 WKKKNIT---TKTKINGSANFSRDEEMKQVNMDSNGLSEYLDLSSGCPVEILTVTE 329
  |:::|
  |:::|
  |:::|
Db 272 DDAKKSIDSLTRIPRIDGDEGRLTKRLTRFODLNQVHGLTVYS-----TVITE 325
  |:::|
  |:::|
  |:::|
QY 330 SVTGISRVVSTNVFFKQHDYILIEFPDYTVLKPSLNFTVYVTRADGNQULIEBRNV 389
  |:::|
  |:::|
  |:::|
Db 326 SGSDMVTEQGGIHIVTSFYQILFYKTPKPYKFGMEYELTVVYVTPDGSPPA-----H 378
  |:::|
  |:::|
  |:::|
QY 390 VITVYQRYVTEWGSNGSNQCKMEAVOKINYTV---QSGTFKIEPFIILDESSELQKAY 446
  |:::|
  |:::|
  |:::|
Db 379 VPVSEALHSE-----CTTISDGAKLILMTPLIQS-----LPIVRNTHGULPRE 425
  |:::|
  |:::|
  |:::|
QY 447 FLGSHSMAVHSLFKSPKTYIQLKTRDENIKVS--PELVVSGN---KRLKELSYW 500
  |:::|
  |:::|
  |:::|
Db 426 ROAKSMATATAVQTOGGSSENYLHVAITSTEIKPDMLPVNFNVNNGANSLNQIKYTYLI 485
  |:::|
  |:::|
  |:::|
QY 501 VSRGQLVAVK-----QNSTWFSLTPENSWTPKACVITYYYEDDEGELISDYALKIPVQ-- 552
  |:::|
  |:::|
  |:::|

```

```

Db 486 LNKGIKFKVGRQRPDRDQNLVTMNLHITPDLIPSRFVAAYQGNNEIVADSWVNDVKDT 545
Qy 553 ---LVFK-----NKIKLYWSKVKAEPSEKSVIRISVTPDSTIGIVADKSVNLNNA5 602
Db 546 CMGTLVYKAGSSRDDRIG-----KPGAAMKIKLE-GDGRARVGLVADVKAAYVLNDK 596
Qy 603 NDIMNEVNHLELXNTGYVLGMFMNSFAVQEGMLVLDLANLT 647
Db 597 YKISOAKIWDITKESDPCGTAGSGQNNLGVFEDGALTLTSTYLNNTYKORSAAKCPQPNR 656
Qy 648 -----KDIYGVYN-----AEYERMEENEG----- 670
Db 657 RRRSSVLLDSKASKAQAQFODQGRKCCBQGMENHPMGYICEKRAKTIQEGDACAFALE 716
Qy 671 --HIV-----DHDPSLGSSPHV-RKHPEETIWL-----DTNMG 702
Db 717 CCHYIKGIRDENQRESELFLARSDFEDELFGDDNIISRDPFESWLTEELTGEPPNQG 776
Qy 703 YRIQEEFVYV---DSITSWATGAVISSEDGLGTT-----PVELQAQPFPIF 751
Db 777 IS-----SKTVPFYLRSITTT-----ELAVGISPTKGICVABPEYIITWKDFID 823
Qy 752 LMLPSYIRGEFALETIIFNYLKDATBEVYKIEKSPFIOLMTSEINATGHQOULVP 811
Db 824 LRLPSYVAKNEQVETRALINYNADEDIYKVELIYNAF-----CSASTEGORHQQPFPIK 879
Qy 812 SEDGATVLPFIRPTHLG--ELPIVTALSPASDAVTOMLVYKAGIEKSYOSILDLT 869
Db 880 ALSRAVPFVIVPLEQGLHDEVYIASRGELASGVAKKLKVPPEGRRKNIVITIELDPS 939
Qy 870 DNRLOST--LKTLSFSPFNTVTGSEVQITAIQD---VLGPSING--LASLIRMPYGC 921
Db 940 VKGVGGTOELVIANKLDDKVPDETEVETRISVLGDPVQAQIENSIDSKNLHLITPESGC 999
Qy 922 GEOMNINPAPI---YLLDYLTKKKOITDMLKEKALSFMNGQYORELLYQREDSFAPG 978
Db 1000 GEOMNITMTBSVIAITYLDTATGQEMENGVDPRETAIQOITGTAQOQNVYKKAHDSVIAFT 1059
Qy 979 NYDSSGSTWLSAFVLRCFLEADPYI-DIDONVLRITWTW--KGHQKNGEWFMDGRYIHS 1036
Db 1060 N--PASSSWLTAIVYVKYLAWSNNVKDISHEIICGVKWLILNRQGDGVYKENAPYIHG 1117
Qy 1037 ELQGNKS---PVTLYAVITSLLGYRKYOPNIDVOESIHFLSEFSRG-----I 1083
Db 1118 EMLCGTGABEASLTFAIVTALLESRSV-----CKEQINILDSINKATDYLLKYEKL 1172
Qy 1084 SDNTLALITVALLSVSPKAKELNMLYTRABEGEMQWVWSESSEKLSMSWORSIDIE 1143
Db 1173 QRPYTLALTAALA-----AABLNDRVLMAASTRRNW-----BEYNARTHNIE 1218
Qy 1144 VAAVALISHFLQFOTSEGIPIRMWLSQRNSLGGFASFTOPTVYALKALSEFAALMNTER- 1202
Db 1219 GTSJALLALALKMKFABVGPVVKLIDQKYYGGTQATVWVYQALABEYIOMPHQD 1278
Qy 1203 TNICVTVTGSSPSR-----LAVQPMAY-----NISANGFPAICQLNIVYVNVK 1247
Db 1279 LMLDISIKLPERVEBERYSINDRNAVQARTVETKLEDFYVSASGDGKATMTILTYVNAQ 1338
Qy 1248 ASGSSRRRRRSIQNOEAPDLVAV-----KENKDILNHYDLVNCISFSGPGRSGMALME 1300
Db 1339 L-----REDANVCNK--PHLDVSVENVELNKQAQKGAARLKLCTRYLLEBVSJTWMIID 1392
Qy 1301 VNLSGFMVPSBEAL-----SLSEFVKKVVEYDH-----GKLNLYDSVNEQO--PCVNTPAV 1349
Db 1393 ISMLTGTFFPDADLKLKLSNGVDRIKSFELIDNMAQKGTYVYIYLDKXSHSEDECLHKIHX 1452
Qy 1350 RNFKVSNTOASVSIIVDYEBRROAVRSYNSEVKJLSSCDLCSVDQGRPCEDGASGSHH 1409
Db 1453 KHFEVGFILQPSKVSVYTNLDEQCTKFHPDKETGLINKICHGNIICRCABETCSLNLQ 1512
Qy 1410 -----SVYIFIFCFKL-----YFMEL 1426
Db 1513 KKIDLOLRIOKACANQVNDVYVYKTLARIBERKGDNDIYFMQV 1553

```

RESULT 14  
 CO3\_RAT STANDARD: PRT: 1663 AA.  
 ID\_CO3\_RAT  
 AC P01026;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-AUG-1990 (Rel. 15, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Complement C3 precursor [Contains: C3a anaphylatoxin].  
 GN C3.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Wistar; TISSUE=Liver;  
 RX MEDLINE=90245672; PubMed=2336397;  
 RA Misumi Y., Sohma M., Ikehara Y.;  
 RT "Nucleotide and deduced amino acid sequence of rat complement C3,"  
 RL Nucleic Acids Res. 18:2178-2178(1990).  
 RN [2]  
 RP SEQUENCE OF 671-748.  
 RX MEDLINE=79062262; PubMed=309768;  
 RA Jacobs J.W., Rubin J.S., Hugli T.E., Bogardt R.A., Mariz I.K.,  
 RA Daniels J.S., Daughaday W.H., Bradshaw R.A.;  
 RT "Purification, characterization, and amino acid sequence of rat  
 anaphylatoxin (C3a)."  
 RL Biochemistry 17:5031-5038(1978).  
 RN [3]  
 RP SEQUENCE OF 1316-1595 FROM N.A.  
 RX MEDLINE=89380332; PubMed=2674144;  
 RA Sundercom S.A., Komm B.S., Ponce-De-Leon H., Yi Z., Teuscher C.,  
 RA Lyette C.R.;  
 RT "Regulation of tissue-specific expression of complement C3,"  
 RL J. Biol. Chem. 264:16941-16947(1989).  
 CC -1- FUNCTION: C3 PLAYS A CENTRAL ROLE IN THE ACTIVATION OF THE  
 COMPLEMENT SYSTEM. ITS PROCESSING BY C3 CONVERTASE IS THE CENTRAL  
 REACTION IN BOTH CLASSICAL AND ALTERNATIVE COMPLEMENT PATHWAYS.  
 CC AFTER ACTIVATION C3B CAN BIND COVALENTLY, VIA ITS RECEPTIVE  
 THIOLESTER, TO CELL SURFACE CARBOHYDRATES OR IMMUNE AGGREGATES.  
 CC -1- FUNCTION: DERIVED FROM PROTEOLYTIC DEGRADATION OF COMPLEMENT C3,  
 C3A ANAPHYLATOXIN IS A MEDIATOR OF LOCAL INFLAMMATORY PROCESS. IT  
 INDUCES THE CONTRACTION OF SMOOTH MUSCLE, INCREASES VASCULAR  
 PERMEABILITY AND CAUSES HISTAMINE RELEASE FROM MAST CELLS AND  
 BASOPHILIC LEUKOCYTES.  
 CC -1- SUBUNIT: C3 PRECURSOR IS FIRST PROCESSED BY THE REMOVAL OF 4 ARG  
 RESIDUES, FORMING TWO CHAINS, BETA & ALPHA, LINKED BY A DISULFIDE  
 BOND. C3 CONVERTASE ACTIVATES C3 BY CLEAVING THE ALPHA CHAIN,  
 RELEASING C3A ANAPHYLATOXIN & GENERATING C3B (BETA CHAIN + ALPHA'  
 CHAIN).  
 CC -1- SIMILARITY: TO C4, C5 AND ALPHA-2-MACROGLOBULIN.  
 CC -1- SIMILARITY: Contains 1 anaphylatoxin-like domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.isb.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X52477; CA36716.1; -;  
 DR EMBL; M29866; AAA40837.1; ALT\_SEQ.  
 DR PIR; S15764; C3RT.  
 DR PDB; 1QOF; 31-JUL-00.  
 DR PDB; 1QSF; 31-JUL-00.  
 DR InterPro; IPR002890; A2M\_N.  
 DR InterPro; IPR000020; Anaphylatoxin.  
 DR InterPro; IPR001599; Macroglobulin2.  
 DR InterPro; IPR001134; Netrin\_C.  
 DR Pfam; PF00207; A2M; 1.

DR Pfam; PF01835; A2M\_N; 1.  
 DR Pfam; PF01821; ANATO; 1.  
 DR Pfam; PF01759; NTR; 1.  
 DR ProDom; PD003264; Anaphylatoxin; 1.  
 DR SMART; SM00104; ANATO; 1.  
 DR SMART; SM00643; C345C; 1.  
 DR PROSITE; PS00477; ALPHA\_2\_MACROGLOBULIN; 1.  
 DR PROSITE; PS01177; ANAPHYLATOXIN\_1; 1.  
 DR PROSITE; PS01178; ANAPHYLATOXIN\_2; 1.  
 KW Complement pathway: Complement alternate pathway: Plasma;  
 KW Inflammatory response; Glycoprotein; Signal; 3D-structure;  
 KW Thioester bond.  
 FT SIGNAL 1 24  
 FT CHAIN 25 1663 COMPLEMENT C3.  
 FT CHAIN 25 666 BETA CHAIN.  
 FT CHAIN 671 1663 ALPHA CHAIN.  
 FT PEPTIDE 671 748 C3A ANAPHYLATOXIN.  
 FT CHAIN 749 1663 C3B (ALPHA' CHAIN).  
 FT SITE 748 749 CLEAVAGE (BY C3 CONVERTASE).  
 FT DOMAIN 693 728 ANAPHYLATOXIN-LIKE.  
 FT DISULFID 558 816 INTERCHAIN (BY SIMILARITY).  
 FT DISULFID 626 661 BY SIMILARITY.  
 FT DISULFID 693 720 BY SIMILARITY.  
 FT DISULFID 694 727 BY SIMILARITY.  
 FT DISULFID 707 728 BY SIMILARITY.  
 FT DISULFID 873 1513 BY SIMILARITY.  
 FT DISULFID 1101 1158 BY SIMILARITY.  
 FT DISULFID 1358 1489 BY SIMILARITY.  
 FT DISULFID 1389 1458 BY SIMILARITY.  
 FT DISULFID 1506 1511 BY SIMILARITY.  
 FT DISULFID 1518 1590 BY SIMILARITY.  
 FT DISULFID 1537 1661 BY SIMILARITY.  
 FT CROSSLINK 1010 1013 Iso-glutamy1 cysteine thioester (Cys-Gln).  
 FT CARBOHYD 939 939 N-LINKED (GLCNAC. . .) (PROBABLE).  
 FT CARBOHYD 1617 1617 N-LINKED (GLCNAC. . .) (PROBABLE).  
 FT CONFLICT 721 722 LK -> KL (IN REF. 2).  
 SQ SEQUENCE 1663 AA; 186460 MW; 2887CCB143CDD4BC CRC64;  
 Query Match 13.1%; Score 966; DB 1; Length 1663;  
 Best Local Similarity 24.3%; Pred. No. 6.4e-45; Indels 326; Gaps 68;  
 Matches 391; Conservative 284; Mismatches 609;  
 QY 3 GPTLTAAHLLCYCTALAAVAPGRPLVYAPGIIIRRGNTVIGVELLE--HCPGQVTK 59  
 DB 6 GSQLV--LLILLASSLLALGSPMVSIIIPNVLRIESETFILAHADAGDVPVTVQ 62  
 QY 60 AELIKTASNTLVSYLAEVGFEGKSEFKYTLPLSLNADLEYELRVGTODEILFSN 118  
 DB 63 DFLKKQVLTSEKTVLTGATGHLNRVFIK--IPASKEFNADKHKVTVVANFGATVEK 119  
 QY 119 STLSPEPKRIISVFIOTDKALYKPKQEVKRIYTLPSDFPKYPTSINILIKDKPSNLIQ 178  
 DB 120 AVLVSPQSGYL--FIQTDKITYPGSTVFRIFTVDNLLVSKTVIVILETDGVPILK 177  
 QY 179 -WLSQSDLDGVISFTQLSHPILGMSIQV--QVNDQTYGSGFVSEVYLPKFEVTLQ 234  
 DB 178 DILSSHQVGIILPLSNIPELVMGMKIRAFEHAPKQFSAFEVKEVLSFEVLVE 237  
 QY 235 --TPLYCSNMSKILNGTTAKTYGKPVKGDTLTF-----LPLSFMGKKNIITK 282  
 DB 238 PTEKFYIIHPPKILEVSIPTARFLYGRNVGTAFVIGVODEDKISIALS-----LTR 290  
 QY 283 TPKINGSANSPFDEMKVNDSSNGISEYLDSSPGPVELITLVESVIGISRNSTNV 342  
 DB 291 VLIEDSGEAVLS--RKVLMDG-----VRPSP-----BALGKSLYVSIVY 330  
 QY 343 FFKQ-----HDYIIFFDYTYTLKPSLNFATATVATVTRADGNQLTEERR 386  
 DB 331 ILSGSDMVAERSGIPYIVSPYQIHFTKPKFKFAMPDVAVFTNPQGS----PAR 386  
 QY 387 NNVTIVTQNNRYEYSGNS-GNQKREAVOKINVTVPQ-----GTFKIEPFLIED 437  
 DB 387 ---VPVTVQ-----GSDAQAALTQDDGVAKLSTVTPNNROPLITIVSTKKEGIPDARQ 435

```

Qy 438 SGE-IQLKAYFLGSSMAVHSLFKSPSTYIQLKTRDENIKVGSFELVSGNRLKEL 496
Db 436 ATRTQAOQPYSTHNSNNYLHL-----SVSRVELKPGD-NLNVNHLRTADGQEAIRYX 489
Qy 497 SYMVTSRGQLVAVK-----ONSTMFSL--TPENSWTPKACIYVY--IEDDG--ELIS 544
Db 490 TYLVNKKGLKLAGQVREPPGQDLVLSLPIPE--FIPSEFLVAYLTIGANGQREVA 547
Qy 545 DYLIKIPVQ-----LVFKKIKLYMSKYVAEPESEKYLRISTQPSIYGIYAVPKSVNL 598
Db 548 DSWMDVDSDSCVGLIVKGDPR---DNQPAHGHTLIRBGNQ-GARVGLVAVDKGVAV 603
Qy 599 MNASNDITMENVHLELYNTG-----YILGEMNNSPAVQO-----ECG- 637
Db 604 LNKKNKLTQSKIMDVEKADICTPGSKGVAGVFMDAGLTPTKNGQLGTQDRDEPECAK 663
Qy 638 -----LWVLTDLNLTQDYLIDGV----- 654
Db 664 PAARRRSVQLMERRMDKAGQYTDKGLKRCCEBGRKIDIPYSCORARLITQGESCLKA 723
Qy 655 -YDNEAVERFMEEN-EGHIV-----DIDFSLGSSPHV-RKHPPEWTW----- 696
Db 724 FMDCCNYITKLREQRHRDHVGLASDVDEDIIPREDIISRHFFPSWMTTEBELKEPK 783
Qy 697 --LDTNMGYRIYQEEFVTVDSITSWVATGVISEDLGLTTPPELQAFQPFIFLNL 754
Db 784 NGISTKV-----NMIFLKDSITTEWELIAYLSLDPKGI-CVADPEIETVMDFFIDLRL 835
Qy 755 PYSVIRGEFALIEITFNLYKATVEKYIE--KSDKPIILMTSEINATGHOQLTVDS 812
Db 836 PYSVIRNEQVEIRAVLFNY-REOEKLYKAVELNHPAFCSMATAK--RYQTEIPIR 890
Qy 813 EDGATVLPFIRPTHG--EIPITVALSPASDAVQMLVYAKGIEKSYOSI-LLD-- 867
Db 891 KSSAVVPYIVLPKLGQEVKAAVFNFIISDVAKILKVPEGRVAKTVAVRTLDE 950
Qy 868 -LTDNRLOS---TLKTLSPFPPTVTGSE--VOITAGVGLPSING--IASIIMP 918
Db 951 HLNQGVQREDVNAADLSQVDP---TDBETRILOQTVAQMAEBAVDEGKHLIYV 1007
Qy 919 YCGGQNMNINFAPNYIYLDYLTKKQOLD--NLKELKALSPRQGYORELLYOREDSFS 975
Db 1008 SGCGQNMNIGMPTVIAYVHLYLQTEQWEKFGLEKQEALELIKKGYTQGLAKFOPISAVA 1067
Qy 976 AFGNVDPSGSTWLSAFVLCFLEADPYIDIDONVLRHTYWL-KGHQKNSGEFMPGRYI 1034
Db 1068 AFNNRPP--STWLTAMWSRFSFLANLAIIDSQVLCGAVKWLILEKQKPDGVFOEDGPYI 1125
Qy 1035 HSELQGG---NKSPTLTAYIVTSLIGYR--KYQONI---DVOESIHFFSEFSRGIS 1084
Db 1126 HQEMIGGRNTEADVSLTAFVIALQEKARDICEQVNSLPISINKAGYILAST-LNQ 1184
Qy 1085 DNYTLALITVAL---SSVGSPPAKALANLTVRAR-OEGMGQFVWSSEKLSDSQPRSL 1140
Db 1185 RPYTVAINGYALALANMKLEBPYLTFLNAXKRNREBERGQQLY----- 1228
Qy 1141 DIEVAAYLLSHFLQFQTSBGIPIRMWLSKQNSLGGFASSTODTYALKALSEFAA-LMN 1199
Db 1229 NVEATSYVALALLLLKDPDSVPVAVRWLDERYYGGGYSTQATFWFALAQYADVPD 1288
Qy 1200 TERTNIQVTVTSPSSPLA-----VVOPMANVISANGFPAICOLANVY 1244
Db 1289 HKDLAMDVSIMHPSSSPVFRLLWESSGLASEETKQEGESLTKKGGQGLTSLVYVY 1348
Qy 1245 NVKASGSSRRRSIONQEAFLDLVAV-----KENKODLNVHVDLVNCTSFSGPGRSGMA 1297
Db 1349 HAKVAGKTCKK-----FDIRVTIKPAPETAKQDAKSSMILDICTRYIGVDATMS 1401
Qy 1298 LMEVNLISGFMPSEALISSEI--VKKVEYDHCKLN-----LYLDSVNET-QFCVNI 1346
Db 1402 ILDISMTMGFIPDTNDELSSGVDRIYSKYEMDKAFSNKNTLIYLEKISHSEBDSLSF 1461

```

```

Qy 1347 PAVNRFKVSNTQDASVSIVDYEPFRRAVRSYNSE-----VKLSSCDLC 1390
Db 1462 KVHGFENVLIOPGSVKYVSYNLEESCTRFYHPEKDDGMLSKLCHNMC 1511
RESULT 15
CO3_MOUSE STANDARD, PRT, 1663 AA.
ID CO3_MOUSE
AC P01027;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Complement C3 precursor (HSE-MSF) [Contains: C3A anaphylatoxin].
GN C3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM LONG).
RX MEDLINE=85038854; PubMed=6208565;
RA Fey G.H., Lundwall A., Wetzel R.A., Tack B.F., de Bruijn M.H.L.,
Domney H.;
RT "Nucleotide sequence of complementary DNA and derived amino acid
sequence of murine complement protein C3."
RL Philos. Trans. R. Soc. Lond., B, Biol. Sci. 306:333-344(1984).
RN [2]
RP SEQUENCE OF 671-1663 FROM N.A. (ISOFORM LONG).
RX MEDLINE=85054819; PubMed=6094532;
RA Wetzel R.A., Lundwall A., Davidson F., Gibson T., Tack B.F., Fey G.H.;
RT "Structure of murine complement component C3. II. Nucleotide sequence
of cloned complementary DNA coding for the alpha chain."
RL J. Biol. Chem. 259:13857-13862(1984).
RN [3]
RP SEQUENCE OF 671-748 FROM N.A.
RX MEDLINE=83117730; PubMed=6961437;
RA Domney H., Wiebauer K., Kazmaler M., Mueller V., Odink K., Fey G.H.;
RT "Characterization of the mRNA and cloned cDNA specifying the third
component of mouse complement."
RL Proc. Natl. Acad. Sci. U.S.A. 79:7619-7623(1982).
RN [4]
RP SEQUENCE OF 658-761 FROM N.A.
RX MEDLINE=84201365; PubMed=6609661;
RA Fey G.H., Wiebauer K., Domney H.;
RT "Amino acid sequences of mouse complement C3 derived from nucleotide
sequences of cloned cDNA."
RL Ann. N.Y. Acad. Sci. 421:307-312(1983).
RN [5]
RP SEQUENCE OF 1-34 FROM N.A.
RX MEDLINE=83117622; PubMed=6985486;
RA Wiebauer K., Domney H., Diggelmann H., Fey G.;
RT "Isolation and analysis of genomic DNA clones encoding the third
component of mouse complement."
RL Proc. Natl. Acad. Sci. U.S.A. 79:7077-7081(1982).
RN [6]
RP SEQUENCE OF 25-41 AND 749-760.
RX MEDLINE=93373334; PubMed=8364938;
RA Hamada U.-I., Gavanagh P.G., Miki K., Nicolson G.L.;
RT "A paracrine migration-stimulating factor for metastatic tumor cells
secreted by mouse hepatic sinusoidal endothelial cells:
identification as complement component C3b."
RL Cancer Res. 53:4418-4423(1993).
RN [7]
RP ALTERNATIVE INITIATION.
RX MEDLINE=95053742; PubMed=7964485;
RA Cahen-Kramer Y., Mattenson I.L., Melchers F.;
RT "The structure of an alternate form of complement C3 that displays
costimulatory growth factor activity for B lymphocytes."
RL J. Exp. Med. 180:2079-2088(1994).
CC -I- FUNCTION: C3 PLAYS A CENTRAL ROLE IN THE ACTIVATION OF THE
CC COMPLEMENT SYSTEM. ITS PROCESSING BY C3 CONVERTASE IS THE CENTRAL
CC REACTION IN BOTH CLASSICAL AND ALTERNATIVE COMPLEMENT PATHWAYS.
CC AFTER ACTIVATION C3B CAN BIND COVALENTLY, VIA ITS REACTIVE

```



CC THIOLESTER, TO CELL SURFACE CARBOHYDRATES OR IMMUNE AGGREGATES.  
 CC -1- FUNCTION: DERIVED FROM PROTEOLYTIC DEGRADATION OF COMPLEMENT C3, C3A ANAPHYLATOXIN IS A MEDIATOR OF LOCAL INFLAMMATORY PROCESS. IT INDUCES THE CONTRACTION OF SMOOTH MUSCLE, INCREASES VASCULAR PERMEABILITY AND CAUSES HISTAMINE RELEASE FROM MAST CELLS AND BASOPHILIC LEUKOCYTES. THE SHORT ISOFORM HAS B-CELL STIMULATORY ACTIVITY.  
 CC -1- SUBUNIT: C3 precursor is first processed by the removal of 4 Arg residues, forming two chains, beta and alpha, linked by a disulfide bond. C3 convertase activates C3 by cleaving the alpha chain, releasing C3a anaphylatoxin and generating C3b (beta chain + alpha chain).  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event-alternative initiation;  
 CC Comment-2 isoforms, long (shown here) and short, are produced by alternative initiation;  
 CC -1- MISCELLANEOUS: C3b IS RAPIDLY SPLIT IN TWO POSITIONS BY FACTOR I AND A COFACTOR TO FORM IC3b (INACTIVATED C3b) AND C3f WHICH IS RELEASED.  
 CC -1- MISCELLANEOUS: IC3b IS THE SLOWLY CLEAVED (POSSIBLY BY FACTOR I) TO FORM C3c AND C3dg. OTHER PROTEASES PRODUCE OTHER FRAGMENTS SUCH AS C3d OR C3g.  
 CC -1- SIMILARITY: TO C4, C5 AND ALPHA-2-MACROGLOBULIN.  
 CC -1- SIMILARITY: Contains 1 anaphylatoxin-like domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.ebi.ac.uk/announcements/](http://www.ebi.ac.uk/announcements) or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC -----  
 DR EMBL; K02782; AAC42013.1; -  
 DR EMBL; J00369; AAA37336.1; -  
 DR EMBL; J00367; AAA37336.1; JOINED.  
 DR EMBL; M33032; AAA37378.1; -  
 DR EMBL; Z37998; CAA86099.2; -  
 DR PIR; A92459; C3MS.  
 DR HSP; P01024; 1C3D.  
 DR MGI; M8227; C3.  
 DR InterPro: IPR002890; A2M\_N.  
 DR InterPro: IPR000020; Anaphylatoxin.  
 DR InterPro: IPR001599; Macroglobulin2.  
 DR InterPro: IPR001134; Netrin\_C.  
 DR Pfam; PF00207; A2M; 1.  
 DR Pfam; PF01835; A2M\_N; 1.  
 DR Pfam; PF01821; ANATO; 1.  
 DR Pfam; PF01759; NTR; 1.  
 DR ProDom; PD003264; Anaphylatoxin; 1.  
 DR SMART; SM00643; C345C; 1.  
 DR PROSITE; PS00477; ALPHA\_2\_MACROGLOBULIN; 1.  
 DR PROSITE; PS01177; ANAPHYLATOXIN\_1; 1.  
 DR PROSITE; PS01178; ANAPHYLATOXIN\_2; 1.  
 KW Complement pathway; Complement alternate pathway; Plasma; Inflammatory response; Glycoprotein; Signal; Alternative initiation; Thioester bond.  
 KW Thioester bond.  
 FT SIGNAL 1 24  
 FT CHAIN 25 1663  
 FT CHAIN 25 1663  
 FT CHAIN 671 1663  
 FT CHAIN 1129 1663  
 FT INIT MET 1129 1129  
 FT PEPTIDE 671 748  
 FT CHAIN 749 1663  
 FT PEPTIDE 749 954  
 FT PEPTIDE 955 1303  
 FT PEPTIDE 955 1001  
 FT PEPTIDE 1002 1303  
 FT PEPTIDE 1304 1320  
 FT SITE 748 749  
 FT SITE 1303 1304

FT SITE 1320 1321  
 FT DOMAIN 693 728  
 FT DISULFID 559 816  
 FT DISULFID 626 661  
 FT DISULFID 693 720  
 FT DISULFID 694 727  
 FT DISULFID 707 728  
 FT DISULFID 873 1513  
 FT DISULFID 1101 1158  
 FT DISULFID 1358 1489  
 FT DISULFID 1389 1458  
 FT DISULFID 1506 1511  
 FT DISULFID 1518 1590  
 FT DISULFID 1537 1661  
 FT DISULFID 1637 1646  
 FT CARBOHYD 939 939  
 FT CARBOHYD 1617 1617  
 FT CARBOHYD 1010 1013  
 SQ SEQUENCE 1663 AA; 186482 MW; DE5546CCT69BEA19 CRC64;  
 (By similarity)

Query Match 13.0%; Score 956; DB 1; Length 1663;  
 Best Local Similarity 24.0%; Pred. No. 2.3e-44;  
 Matches 384; Conservative 286; Mismatches 626; Indels 302; Gaps 66;

QY 3 GPRLLTAHLICVCTAALAAVAPGRFLVAPGIRPGANTIGVLELE--HCPQYTVK 59  
 DB 6 GSQULVLLILLASSPLALGI---PMYSIIPNVLRLESEITVLEHNDAGDIPVTVQ 62  
 QY 60 AELLKTASNLTVSLVLE-AEGVPEKSEFVTLTP-SLPNSADEIYEIVRTGRQDEILFS 117  
 DB 63 DFLKRVLTSEKVTITGASGHLRSVSIK--IPASKEFNSDKGHKVVAVANFGETIVE 119  
 QY 118 NSTRLESETKRISVFIOTDKALKPKOEKFRIVTLFSDKPKYKTSINTLINDPKSLIQ 177  
 DB 120 KAVWVSFQSGYL--FIQTDTITTPSGTVLYRIFVDNNLLIPGKTVVILIEFPGIPVK 177  
 QY 178 Q-WLSQSDLGIVISKTFQLSHPIIDGMSIQV---QVNDQTVYVSFQVSEVVLPRFEVTL 233  
 DB 178 RDLISSNNGGILPLSNIPELVNMGMKIRAYEHAKPIFAFEFVKYVLPSEFVRV 237  
 QY 234 Q--TPLYCSNMSKRLNGITAKTYTGKPKVGDVTLFPLSPWKKK---NITKTFKI 286  
 DB 238 EPTETFYIIDDPNGLEVISIATAFLYGKVDGTFVIFGVOD--GDKKISLAHSITRVVIE 295  
 QY 287 NGSANFEFNDKEMKAWNDSEN-----GLSEYLDLSSGPEVILITVETSTGISRNSTN 341  
 DB 296 DGVGDVAVLTRKVLMEGVRPSNADALVKSILYVS-----VTVILHSGSDMVEARSG 346  
 QY 342 VFFKQHDYIEFPDYTVLKRPSINFATVATVTRADGNQLTLEBRNNVITVQRYNTEY 401  
 DB 347 IPIVTSYQIHFKTPFFKPPMPFDMVFTVNPDSPPA-----SKLVVTD----- 393  
 QY 402 WSGSNS-GNGKMEAVOKINTVPOG-----TKIEPFIIDSESE---LQKAYFLGSK 451  
 DB 394 --GSNAKALIQDDGVAKLSINTPNSRQPLTITVTRKTDLPESRQATKMEAPHYSTMAN 451  
 QY 452 SSMAVHSLFSPKTYIQLTRDENIVKSPFELVYSGNRLKELSTMVVSRQLVAVGK 511  
 DB 452 SNNYLIH----SVSRMELKPGD-NLVNPHLRTDPGHEAKIRYTYLVNKKGLVAGR 505  
 QY 512 -----ONSTMFSL--TPENSWTPKACVIVY-----IEDGELISDVLKIPVQ-----L 553  
 DB 506 QVREPGQDLVLSLPIFE--FIPSPFLVAVYTLIGASGQREVAVDSVWDVADSCIGTL 563  
 QY 554 VFNKKIKLWYSKYABSEKVSRLISVTPDSIVGIYAVDKSVNLNMAASNDITMENVHE 613  
 DB 564 VVKGDPR-----DNHLARQQTTLRIEIGNQ--GARVGLVAVDKGVFLNKKKLTQSKIMDV 618  
 QY 614 LELVNTG-----YIIGMFMNSPAVFQ-----EC----- 636  
 DB 619 VEKADICTPDSGKNVAGVMDAGLAKTISQGLQTEORADLECTKPAARRRRSVOLMERR 678

[illegible]

**THIS PAGE BLANK (b)(7)(D)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2004, 18:09:19 ; Search time 53 Seconds  
(without alignments)  
6952.814 Million cell updates/sec

Title: US-10-020-095-4

Perfect score: 7348

Sequence: 1 MGGPPLTAHLICVCTAL.....HSSVIFPCFKLYEMELWL 1428

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL\_23.\*

1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7326.5	99.7	1445	Q8TJD3	Q8TJD3 homo sapien
2	5360	72.9	1442	Q8T422	Q8T422 mus musculu
3	4305.5	58.6	854	Q8N3A7	Q8N3A7 homo sapien
4	3360	45.7	665	Q8N915	Q8N915 homo sapien
5	1850	25.2	1519	Q46015	Q46015 caenorhabdi
6	1779.5	24.2	1809	Q8T398	Q8T398 ciona intes
7	1703	23.2	1388	Q8IPHS	Q8IPHS drosophila
8	1699.5	23.1	1397	Q9VLZ0	Q9VLZ0 drosophila
9	1693	23.0	1408	Q8IPH4	Q8IPH4 drosophila
10	1692	23.0	1420	Q9NFW7	Q9NFW7 drosophila
11	1688.5	22.3	1399	Q9VLV9	Q9VLV9 drosophila
12	1638	22.3	1507	Q01717	Q01717 limulus sp.
13	1542.5	21.0	1884	Q9TUD7	Q9TUD7 homo sapien
14	1542.5	21.0	1885	Q81ZJ3	Q81ZJ3 homo sapien
15	1491	20.3	1476	Q60486	Q60486 cavia porce
16	1463.5	19.9	1496	Q9NFW5	Q9NFW5 drosophila

17	1461.5	19.9	1496	5	Q9VIT9	Q9VIT9 drosophila
18	1460	19.9	1354	5	Q9NFW8	Q9NFW8 drosophila
19	1455.5	19.8	1494	5	Q8T776	Q8T776 onthodora
20	1428.5	19.4	1732	5	Q969A4	Q969A4 branchiosto
21	1425.5	19.4	1469	5	Q9NFW6	Q9NFW6 drosophila
22	1425.5	19.4	1469	5	Q9NFW6	Q9NFW6 drosophila
23	1424	19.4	1503	13	Q91076	Q91076 lampetra ja
24	1417	19.3	1487	11	Q03626	Q03626 rattus norv
25	1384	18.8	1461	13	Q12978	Q12978 xenopus lae
26	1384	18.8	1500	11	Q63041	Q63041 rattus norv
27	1384	18.8	1500	11	Q63332	Q63332 rattus norv
28	1374.5	18.7	1464	11	Q60488	Q60488 cavia porce
29	1354	18.4	1269	5	Q9NKT75	Q9NKT75 drosophila
30	1328.5	18.1	1760	5	Q9VLT3	Q9VLT3 drosophila
31	1321	18.0	1442	13	Q9PVU5	Q9PVU5 cyprinus ca
32	1310.5	17.8	1728	5	Q817P1	Q817P1 swiffla exs
33	1284.5	17.5	1406	13	Q9PVU4	Q9PVU4 cyprinus ca
34	1233.5	16.8	1340	5	Q9GYW4	Q9GYW4 anopheles g
35	1105.5	15.0	1699	5	Q44344	Q44344 strongyloce
36	1099.5	15.0	798	5	Q8WTP94	Q8WTP94 drosophila
37	1079.5	14.7	1683	13	Q91741	Q91741 xenopus lae
38	1073	14.6	1700	13	Q91933	Q91933 cyprinus ca
39	1054.5	14.4	1489	13	Q8QSD4	Q8QSD4 rana catesb
40	1054	14.3	1716	13	Q91932	Q91932 cyprinus ca
41	990	13.5	1652	13	Q90633	Q90633 gallus gall
42	985	13.4	1738	11	Q70346	Q70346 mus musculu
43	983	13.4	1614	13	Q98977	Q98977 oncothynhu
44	965	13.1	1662	13	Q98TS6	Q98TS6 anathchas
45	963	13.1	785	11	Q62591	Q62591 rattus norv

## ALIGNMENTS

ID	Q8TJD3	PRELIMINARY;	PRT; 1445 AA.
AC	Q8TJD3		
DT	01-JUN-2002 (TREMBLrel. 21, Created)		
DT	01-JUN-2002 (TREMBLrel. 21, Last sequence update)		
DT	01-OCT-2002 (TREMBLrel. 22, Last annotation update)		
DE	CD109.		
GN	CD109.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=21849742; PubMed=11861284.		
RA	Lin M.F., Sutherland D.R., Horsfall W., Totty N., Yeo E., Nayar R.,		
RA	Wu X.F., Schun A.C.;		
RT	"Cell surface antigen CD109 is a novel member of the alpha(2)		
RT	macroglobulin/C3, C4, C5 family of thioester-containing proteins."		
RL	Blood 99:1683-1691(2002).		
DR	EMBL: AF410459; AAL84159.1; -		
DR	InterPro: IPR002890; A2M_N.		
DR	InterPro: IPR01599; Macrogloblna2.		
DR	Pfam: PF00207; A2M; 1.		
DR	Pfam: PF01835; A2M_N; 1.		
DR	PROSITE: PS00477; ALPhA.2 MACROGLOBULIN; 1.		
SO	SEQUENCE 1445 AA; 161719 MW; 6EBED2D936AF310D CRC64;		
Qy	Query Match	99.7%; Score 7326.5; DB 4; Length 1445;	
Qy	Best Local Similarity	98.8%; Pred. No. 0;	
Qy	Matches 1427; Conservative	1; Mismatches	0; Indels 17; Gaps 1;
Db	1 MGGPPLTAHLICVCTALAVAPGPRVTAAGIRGAVNTTGVLEHCPGQVYTKA 60		
Qy	61 ELKTAASLITVSVLEAGVFKEGSKTLTLPSPNSADEIYELRVTRTDEILFNSST 120		

```

Db 61 ELKTSANLTVSLAEVGEKGSFKTLPLSLPLNSADEIYELRVGTQDEILFNSNT 120
QY 121 RLSEFKRIISVFIQTDKALYKPKQEVKFRIVTLFSDPEKPYKTSNLIILKDPKSNLIQOML 180
Db 121 RLSEFKRIISVFIQTDKALYKPKQEVKFRIVTLFSDPEKPYKTSNLIILKDPKSNLIQOML 180
QY 181 SQOSDLGVISKTFOQLSSHPILGDMISIQOVNDQTYQSPQVSEYVLPKPEVTLQTPLYCS 240
Db 181 SQOSDLGVISKTFOQLSSHPILGDMISIQOVNDQTYQSPQVSEYVLPKPEVTLQTPLYCS 240
QY 241 MNSKHLNGITTAKTGYKPKVGDVTLTFLPLSPWGKKKNTTKTKINGSNANFSNDEBMK 300
Db 241 MNSKHLNGITTAKTGYKPKVGDVTLTFLPLSPWGKKKNTTKTKINGSNANFSNDEBMK 300
QY 301 NVMSSNGLSEYLDLSSPGVEILTTVESVTGSRNVSTNVFKQHDYIIIEFDYTVL 360
Db 301 NVMSSNGLSEYLDLSSPGVEILTTVESVTGSRNVSTNVFKQHDYIIIEFDYTVL 360
QY 361 KPSLNFATVYKTRADGNQTLLEBRNNVITVTQRYNTEYWSGNSGNQMEAVQKINY 420
Db 361 KPSLNFATVYKTRADGNQTLLEBRNNVITVTQRYNTEYWSGNSGNQMEAVQKINY 420
QY 421 TVPQSGTFKLEPFLLEDSSELQKAYFLGSKSMAVHSLPSPKTYIOLKTRDENIKYG 480
Db 421 TVPQSGTFKLEPFLLEDSSELQKAYFLGSKSMAVHSLPSPKTYIOLKTRDENIKYG 480
QY 481 SPFLVYVSGNKLKELSYMVVSRGOLAVGKONSTWELFPENSWTPKACIYIYIEDDG 540
Db 481 SPFLVYVSGNKLKELSYMVVSRGOLAVGKONSTWELFPENSWTPKACIYIYIEDDG 540
QY 541 EIIISDLKIPVQVLFKNKIKLYMSKVAEPSEKYSLSISVTPDPSIYIGIYAVDKSVLNM 600
Db 541 EIIISDLKIPVQVLFKNKIKLYMSKVAEPSEKYSLSISVTPDPSIYIGIYAVDKSVLNM 600
QY 601 ASNDITMENVHHELVNTGYLLGMPNNSPVPFECGLMUTDNLTKDYIDGYDIAEX 660
Db 601 ASNDITMENVHHELVNTGYLLGMPNNSPVPFECGLMUTDNLTKDYIDGYDIAEX 660
QY 661 AERFMEENEGHIVIHDFSLGSSPHVKGHPETIMWIDJNMGYIYOEFVTVDSITSW 720
Db 661 AERFMEENEGHIVIHDFSLGSSPHVKGHPETIMWIDJNMGYIYOEFVTVDSITSW 720
QY 721 VATGFVISEDGLGLTTPVLEQAPQPFIFLNLPSYIRGEPEALETIFNYLKADATEV 780
Db 721 VATGFVISEDGLGLTTPVLEQAPQPFIFLNLPSYIRGEPEALETIFNYLKADATEV 780
QY 781 KVIIEKSDKPDIIIMTSNEINATGHQOTLLVSEEDGATVLEPIRPHLGEIPIYVLTALSP 840
Db 781 KVIIEKSDKPDIIIMTSNEINATGHQOTLLVSEEDGATVLEPIRPHLGEIPIYVLTALSP 840
QY 841 ASDAVTOMILVKGIEKSYQSILDLTDNRLOSTLKTLSFSPPTVTVGSEKVOITAI 900
Db 841 ASDAVTOMILVKGIEKSYQSILDLTDNRLOSTLKTLSFSPPTVTVGSEKVOITAI 900
QY 901 GDVLGPIINGLASLIRMPYGGCGEONMIFAPNIYILDYLTKKQJLTNLEKALSFMRQ 960
Db 901 GDVLGPIINGLASLIRMPYGGCGEONMIFAPNIYILDYLTKKQJLTNLEKALSFMRQ 960
QY 961 YQRELLVQREDESSAGANDPSSGTWLSAFVLCFLEADPYIDIDONVLRHYTWLKG 1020
Db 961 YQRELLVQREDESSAGANDPSSGTWLSAFVLCFLEADPYIDIDONVLRHYTWLKG 1020
QY 1021 OKSNGEPMDEGRVHSELOGNKSPTVLTATVSLGGRYQONIDVQESIHFLSEFS 1080
Db 1021 OKSNGEPMDEGRVHSELOGNKSPTVLTATVSLGGRYQONIDVQESIHFLSEFS 1080
QY 1081 RGISDNVTALITVLSVSGPKAKALNMLTWABEGEGMGQFVSSSEKLSDSWQPSRL 1140
Db 1081 RGISDNVTALITVLSVSGPKAKALNMLTWABEGEGMGQFVSSSEKLSDSWQPSRL 1140
QY 1141 DIEVAAYALSHFQFOTSGEPIPMRWLSRORNSLGFPASTQDTTVALKALSEFALMNT 1200
Db 1141 DIEVAAYALSHFQFOTSGEPIPMRWLSRORNSLGFPASTQDTTVALKALSEFALMNT 1200

```

```

QY 1201 ERTNIQVTVGSPSPSP-----LAVOPMAVNISANGFALICOLNVV 1243
Db 1201 ERTNIQVTVGSPSPSPKFLIDTNRLLQTAELAVOPMAVNISANGFALICOLNVV 1260
QY 1244 YNVKAGSSRRRRSIONOEAFLDVAVKENKDDLNVHVDLVCTSESGPGRSGMALMEVNL 1303
Db 1261 YNVKAGSSRRRRSIONOEAFLDVAVKENKDDLNVHVDLVCTSESGPGRSGMALMEVNL 1320
QY 1304 LSGFMVPSKISLSETVKVEYDHGKLNLYLDSVNETOPCVNI PAVRNPKVSTODASYS 1363
Db 1321 LSGFMVPSKISLSETVKVEYDHGKLNLYLDSVNETOPCVNI PAVRNPKVSTODASYS 1380
QY 1364 IVDYEPFRQAVASYNSEVWLSGCDLCSVOGCRPCEDGASGSHHSYIIFCFLLYE 1423
Db 1381 IVDYEPFRQAVASYNSEVWLSGCDLCSVOGCRPCEDGASGSHHSYIIFCFLLYE 1440
QY 1424 MELML 1428
Db 1441 MELML 1445

RESULT 2
QY 08R422 PRELIMINARY; PRT; 1442 AA.
ID 08R422;
AC 08R422;
DT 01-JUN-2002 (TREMBLREL. 21, Created)
DT 01-JUN-2002 (TREMBLREL. 21, Last sequence update)
DE 01-OCT-2002 (TREMBLREL. 22, Last annotation update)
DE GPI-anchored alpha-2 macroglobulin-related protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Hashimoto M., Ichihara M., Takahashi M.;
RT "Cloning and characterization of GPI-anchored alpha-2 macroglobulin-
RT related protein.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY083458; AAM00021.1; -.
DR InterPro: IPR002890; A2M_N.
DR Pfam: PF00207; A2M; 1.
DR Pfam: PF01835; A2M_N; 1.
DR PROSITE: PS00477; ALPHA 2 MACROGLOBULIN; 1.
SQ SEQUENCE 1442 AA; 161658 MW; E2B9671BB8A5BAA7 CRC64;

Query Match 72.9%; Score 5360; DB 11; Length 1442;
Best Local Similarity 71.6%; Pred. No. 4.1e-318;
Matches 1038; Conservative 176; Mismatches 202; Indels 34; Gaps 8;

QY 1 MGPSPILTAHLICVCTALAAVAPGRPLVTAGTIRRGNAVITIGVLEHCHPSQTVVA 60
Db 1 MGRSRLISAHLICLCAVLA-APGSRFLVTAGTIRPGNAVITIGVLDLNSPPQVLVA 59
QY 61 ELKTSANLTVSLAEVGEKGSFKTLPLSLPLNSADEIYELRVGTQDEILFNSNT 120
Db 60 QVTKINSKRSRLIEAGVFRHGHFTLVLPALPLSSADKIYELHNGSSENIIVSNRT 119
QY 121 RLSEFKRIISVFIQTDKALYKPKQEVKFRIVTLFSDPEKPYKTSNLIILKDPKSNLIQOML 180
Db 120 RLFEESKISIVLIQTDKAFKPKQEVKFRIVTLFSDPEKPYKTSNLIILKDPKSNLIQOML 179
QY 181 SQOSDLGVISKTFOQLSSHPILGDMISIQOVNDQTYQSPQVSEYVLPKPEVTLQTPLYCS 240
Db 180 SQGDGADVSTFQJSSNPILFGDMSIQOVNDQTYQSPQVSEYVLPKPEVTLQTPLYCS 239
QY 241 MNSKHLNGITTAKTGYKPKVGDVTLTFLPLSPWGKKKNTTKTKINGSNANFSNDEBMK 300
Db 240 LNSKQUNGVIAKTYTGKPEVKSLSLPLSPWGKKKNTTKTKINGSNANFSNDEBMK 299
QY 301 NVM-----SSNGLSEYLDLSSPGVEILTTVESVTGSRNVSTNVFKQHDYIIIEFDY 354

```

Db 300 KWNMLKPLTDBSEGSYENDPSPFGPAKELIATVTSLETSISRMASINVFCKHDYIELF 359  
 Qy 355 DYTTLVKPSLNTFAATVYTRADONOLTEBERRNVITVYTOR--NYTEWSSGNSGNOM 412  
 Db 360 DYTTLVKPSLNTFAATVYTRADONOLTEBERRNVITVYTOR--NYTEWSSGNSGNOM 414  
 Qy 413 EAVOKINTVPOSGTEKIEPILDESSLOLKAAYFLGSKSMAVHSLFKSPKTYIQLNT 472  
 Db 415 DYTTLVKPSLNTFAATVYTRADONOLTEBERRNVITVYTOR--NYTEWSSGNSGNOM 474  
 Qy 473 RBNNTKVSFPFELVYSGNKLKELSTMYVSRQOLVAVGQNSTMSLTPEBNTWTPACVY 532  
 Db 475 RBNNTKVSFPFELVYSGNKLKELSTMYVSRQOLVAVGQNSTMSLTPEBNTWTPACVY 534  
 Qy 533 VYIEDDGIISDVLTIPQVLFKPKIKLYMSKVKAPESEKSLRISVYOPSIVYIVAV 592  
 Db 535 AYIYAEDEGIINDILKIPQVLFKPKIKLYMSKVKAPESEKSLRISVYOPSIVYIVAV 594  
 Qy 593 DKSVMIMNANDITMENVHELELYNTGYLLGMFNSFAVPOECGLMVLTDANLTQYID 652  
 Db 595 DKSVMIMNANDITMENVHELELYNTGYLLGMFNSFAVPOECGLMVLTDANLTQYID 654  
 Qy 653 GYDVAEYAKERMEENEGHIVDIDFSLGSSPHVRKHPEPTWIMDTNNGYRIYOEFEV 712  
 Db 655 EYDTEEYSEKPAEENEGHIVDIDFSLGSSPHVRKHPEPTWIMDTNNGYRIYOEFEV 714  
 Qy 713 VPDSTISWATGVFVSEDGLGLTTTPVLOAPOPFFIPLNLPYSYIRBEEPALETTYN 772  
 Db 715 VPDSTISWATGVFVSEDGLGLTTTPVLOAPOPFFIPLNLPYSYIRBEEPALETTYN 774  
 Qy 773 YLKDAETEVCVILIEKSDKFDILMTSEINATGHQOQLLVSESDGATVLPFRPHLGEIP 832  
 Db 775 YLKDAETEVCVILIEKSDKFDILMTSEINATGHQOQLLVSESDGATVLPFRPHLGEIP 834  
 Qy 833 TTTALSPASDAVYQMLIVKAEIGKYSQSILDLTDNRLOSTLKTLSFSPPTVTGS 892  
 Db 835 TTTALSPASDAVYQMLIVKAEIGKYSQSILDLTDNRLOSTLKTLSFSPPTVTGS 894  
 Qy 893 EKVQITAGIDVIGPISNGLASLIRMPYGCGEQNMIFAPNIYILDVLTCKKQITDNLK 952  
 Db 895 EKVQITAGIDVIGPISNGLASLIRMPYGCGEQNMIFAPNIYILDVLTCKKQITDNLK 954  
 Qy 953 ALSFMRQYORELLIYOREDESFSAFGNYDPGSGTWLSAFVLRCPLEADPYIIDQVLR 1012  
 Db 955 ALSFMRQYORELLIYOREDESFSAFGNYDPGSGTWLSAFVLRCPLEADPYIIDQVLR 1014  
 Qy 1013 TTTMLKQKSGNPFWDPGRVHISELQGNKSPVTLTAYIVTSLGGRKYQPNIDVQSI 1072  
 Db 1015 TTTMLKQKSGNPFWDPGRVHISELQGNKSPVTLTAYIVTSLGGRKYQPNIDVQSI 1074  
 Qy 1073 HFLSEFSRIGISDNTYLLALITVALSVSGPKAKALNMLTWRAEBOGQMFVSSSEKLS 1132  
 Db 1075 HFLSEFSRIGISDNTYLLALITVALSVSGPKAKALNMLTWRAEBOGQMFVSSSEKLS 1134  
 Qy 1133 DSMOPSLIEVAAYVALSHPLQFOTSEGIPIIMRWLSROBNSLGGASQODTIVALKALS 1192  
 Db 1135 DSMOPSLIEVAAYVALSHPLQFOTSEGIPIIMRWLSROBNSLGGASQODTIVALKALS 1193  
 Qy 1193 BEPALMTERINIQTVTGPPSP-----SPLAVQPMANVISANGFG 1235  
 Db 1194 BEPALMTERINIQTVTGPPSP-----SPLAVQPMANVISANGFG 1253  
 Qy 1236 AICQALNVYVKAAGSSRRRSTIQNOEAPDLVAKENKODLNHVDLVNCTSPSGRSG 1295  
 Db 1254 AICQALNVYVKAAGSSRRRSTIQNOEAPDLVAKENKODLNHVDLVNCTSPSGRSG 1312  
 Qy 1296 MALMEVNLISGFVPEASISLSEFVKVYEDHGKLNLYDSVNETOFCNIAVANEKFS 1355  
 Db 1313 MALMEVNLISGFVPEASISLSEFVKVYEDHGKLNLYDSVNETOFCNIAVANEKFS 1372  
 Qy 1356 NTQDASVSIIVDYPERRAQVRSYNSVFKLSCDLCSVOGCRPCEDGASGSHHSVIFI 1415  
 Db 1372 NTQDASVSIIVDYPERRAQVRSYNSVFKLSCDLCSVOGCRPCEDGASGSHHSVIFI 1415

Db 1373 NIRDGSVSMVDYPERRAQVRSYNTQVKLSGCCYLSPT-NCKSHTDGATSLRRSSLLV 1431  
 Qy 1416 FCFKLLYFME 1425  
 Db 1432 FCSVLLYFVQ 1441  
 RESULT 3  
 ID 08N3A7 PRELIMINARY; PRT, 854 AA.  
 AC 08N3A7  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Hypothetical protein (Fragment).  
 GN DKFPP621111.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxId=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Melanoma;  
 RA Blum H., Bauersachs S., Mewes H.W., Well B., Wiemann S.;  
 RL Submitted (JUL-2002) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; AL834478; CAD39137.1; -;  
 DR InterPro; IPR001599; Macroglloblnz.  
 DR Pfam; PF00207; AZM; 1.  
 DR PROSITE; PS00477; ALPHA\_2\_MACROGLOBULIN; 1.  
 KW Hypothetical protein.  
 FT NON TER 1  
 SQ SEQUENCE 854 AA; 95599 MW; 7B51863A2565C004 CRC64;  
 Query Match 58.6%; Score 4305.5; DB 4; Length 854;  
 Best Local Similarity 97.7%; Pred. No. 4.7e-254;  
 Matches 834; Conservative 1; Mismatches 2; Indels 17; Gaps 1;  
 Qy 592 VDKSVNLNANSDITMENVHELELYNTGYLLGMFNSFAVPOECGLMVLTDANLTQYI 651  
 Db 1 VDKSVNLNANSDITMENVHELELYNTGYLLGMFNSFAVPOECGLMVLTDANLTQYI 60  
 Qy 652 DGVDNAEYAKERMEENEGHIVDIDFSLGSSPHVRKHPEPTWIMDTNNGYRIYOEFEV 711  
 Db 61 DGVDNAEYAKERMEENEGHIVDIDFSLGSSPHVRKHPEPTWIMDTNNGYRIYOEFEV 120  
 Qy 712 TTPDSITSWATGVFVSEDGLGLTTTPVLOAPOPFFIPLNLPYSYIRBEEPALETTYN 771  
 Db 121 TTPDSITSWATGVFVSEDGLGLTTTPVLOAPOPFFIPLNLPYSYIRBEEPALETTYN 180  
 Qy 772 NYLKDAETEVCVILIEKSDKFDILMTSEINATGHQOQLLVSESDGATVLPFRPHLGEIP 831  
 Db 181 NYLKDAETEVCVILIEKSDKFDILMTSEINATGHQOQLLVSESDGATVLPFRPHLGEIP 240  
 Qy 832 ITTVALSPASDAVYQMLIVKAEIGKYSQSILDLTDNRLOSTLKTLSFSPPTVTGS 891  
 Db 241 ITTVALSPASDAVYQMLIVKAEIGKYSQSILDLTDNRLOSTLKTLSFSPPTVTGS 300  
 Qy 892 SERVOITAGIDVIGPISNGLASLIRMPYGCGEQNMIFAPNIYILDVLTCKKQITDNLK 951  
 Db 301 SERVOITAGIDVIGPISNGLASLIRMPYGCGEQNMIFAPNIYILDVLTCKKQITDNLK 360  
 Qy 952 KALSMRQYORELLIYOREDESFSAFGNYDPGSGTWLSAFVLRCPLEADPYIIDQVLR 1011  
 Db 361 KALSMRQYORELLIYOREDESFSAFGNYDPGSGTWLSAFVLRCPLEADPYIIDQVLR 420  
 Qy 1012 RYTYMLKQKSGNPFWDPGRVHISELQGNKSPVTLTAYIVTSLGGRKYQPNIDVQSI 1071  
 Db 421 RYTYMLKQKSGNPFWDPGRVHISELQGNKSPVTLTAYIVTSLGGRKYQPNIDVQSI 480  
 Qy 1072 IHFLSEFSRIGISDNTYLLALITVALSVSGPKAKALNMLTWRAEBOGQMFVSSSEKLS 1131  
 Db 481 IHFLSEFSRIGISDNTYLLALITVALSVSGPKAKALNMLTWRAEBOGQMFVSSSEKLS 540

```

QY 1132 SDSWQPSRLDIEVAAYALLSHFLQFQISEGIPIMRWLSRGRNSLGGFASDTDTVALKAL 1191
DB 541 SDSWQPSRLDIEVAAYALLSHFLQFQISEGIPIMRWLSRGRNSLGGFASDTDTVALKAL 600
QY 1192 SEPALMTERNTNQTVTGPPSSPVPKELIDHNRLLLQTAELAVVQPTAVNLSANGFG 1234
DB 601 SEPALMTERNTNQTVTGPPSSPVPKELIDHNRLLLQTAELAVVQPTAVNLSANGFG 660
QY 1235 FAICQNVVNVKASGSSRRRRSIOQOAFDLDAVAKENKDLNHDVLANCTSPSGPRS 1294
DB 661 FAICQNVVNVKASGSSRRRRSIOQOAFDLDAVAKENKDLNHDVLANCTSPSGPRS 720
QY 1295 GMALMEVNLISGFVWPSAISLSETVKKEVYDHGKLNLYDSVNETQPCVNI PAVRNPKV 1354
DB 721 GMALMEVNLISGFVWPSAISLSETVKKEVYDHGKLNLYDSVNETQPCVNI PAVRNPKV 780
QY 1355 SNTQDASVSYDYVEPRQAVRSYKSEVYKSSCDLCSVQCCRCEDGASGSHHSSVIF 1414
DB 781 SNTQDASVSYDYVEPRQAVRSYKSEVYKSSCDLCSVQCCRCEDGASGSHHSSVIF 840
QY 1415 IFCEFLTYFMELML 1428
DB 841 IFCEFLTYFMELML 854

RESULT 4
Q8N915 PRELIMINARY; PRT; 665 AA.
AC Q8N915.
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein FLJ38569.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Oshima A., Takahashi-Fujii A., Tanase T., Imose N., Takeuchi K.,
RA Arita M., Muraashino K., Yuuki H., Hara H., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Magatsuma M.,
RA Murakawa K., Kanahori K., Sugiyama A., Kawakami B., Suzuki Y.,
RA Sugano S., Nagahori K., Maezono Y., Nagai K., Isogai T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK095888; BAC04642.1; -
DR InterPro; IPR002890; A2M_N.
DR Pfam; PF01835; A2M_N; 2.
KM Hypothetical protein.
SQ SEQUENCE 665 AA; 74411 MW; CC73982C8E88647E CRC64;

Query Match 45.7%; Score 3360; DB 4; Length 665;
Best Local Similarity 100.0%; Pred. No. 1.6e-196;
Matches 657; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 241 MNSKHLNGITTAATYTYGPKVGDVTLFLPLSFGKKKNTTKTKINGSANFSFNDSEMK 300
DB 241 MNSKHLNGITTAATYTYGPKVGDVTLFLPLSFGKKKNTTKTKINGSANFSFNDSEMK 300
QY 301 NVMDSSNGLSSEYDLSPGVEVILTTYSVTGISRNVSINVPKQHDYIIIEFDYTVL 360
DB 301 NVMDSSNGLSSEYDLSPGVEVILTTYSVTGISRNVSINVPKQHDYIIIEFDYTVL 360
QY 361 KPSLNFATYKVRADGNQTLERRNNVITYTORNYTEYSGNSGNQKMEAVOKINY 420
DB 361 KPSLNFATYKVRADGNQTLERRNNVITYTORNYTEYSGNSGNQKMEAVOKINY 420
QY 421 TVPQSGTFKIEPFLIEDSSELQKAYFLGSKSSMAVHSLFKSPSKTYIQKTRDENIKVG 480
DB 421 TVPQSGTFKIEPFLIEDSSELQKAYFLGSKSSMAVHSLFKSPSKTYIQKTRDENIKVG 480
QY 481 SPPELVVSGNKRLEKESYMYVSRGQLVANGKNSYMFSLTPENSWPKACVITYYIEDDG 540
DB 481 SPPELVVSGNKRLEKESYMYVSRGQLVANGKNSYMFSLTPENSWPKACVITYYIEDDG 540
QY 541 EITSDVLTPIVOLVFNKIKLYMSKYKAPSEKVSRLISTOPDSTIGVAVDKSVNLN 600
DB 541 EITSDVLTPIVOLVFNKIKLYMSKYKAPSEKVSRLISTOPDSTIGVAVDKSVNLN 600
QY 601 ASNDITMENYVHLELYNTGYLGMPNNSFAVFOEGGLWLTDPANLTQYIDGVYDN 657
DB 601 ASNDITMENYVHLELYNTGYLGMPNNSFAVFOEGGLWLTDPANLTQYIDGVYDN 657

RESULT 5
Q46015 PRELIMINARY; PRT; 1519 AA.
AC Q46015.
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical 168.4 kDa protein ZK337.1.
GN ZK337.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OC NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA White S.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; ZK337.1B (SHOWN HERE) AND
CC ZK337.1A; ARE PRODUCED BY ALTERNATIVE SPLICING.
DR EMBL; Z82090; CAB05007.1; -
DR EMBL; Z82090; CAB05006.1; -
DR HSP; P01024; ICD3.
DR WormPep; ZK337.1a; CE16718.
DR WormPep; ZK337.1b; CE16719.
DR InterPro; IPR000847; A2M_N.
DR InterPro; IPR000847; HTH_LysR.
DR InterPro; IPR001599; Macroglolnab2.
DR InterPro; IPR001917; Nhrtransf_2.
DR Pfam; PF00207; A2M; 1.
DR Pfam; PF01835; A2M_N; 1.
DR PROSITE; PS00599; AA TRANSFER CLASS 2; 1.
DR PROSITE; PS00477; ALPHA 2 MACROGLOBULIN; 1.
DR PROSITE; PS00044; HTH_LYSR FAMILY; 1.
KM Alternative splicing, Hypothetical protein.
FT VARSPLIC 714 724 MISSING (IN ISOFORM ZK337.1A).
SQ SEQUENCE 1519 AA; 168385 MW; 78139ACB564AB742 CRC64;

Query Match 25.2%; Score 1850; DB 5; Length 1519;
Best Local Similarity 31.6%; Pred. No. 8.6e-104;
Matches 477; Conservative 290; Mismatches 563; Indels 180; Gaps 44;

```



Db 28 VSTTAPVKA-TWMLVAVRPPQPSVCMNLKQATDEDMIVRIEVERTNETIAR 86  
 QY 75 EAEVPEKSGFKTLTLPSPILNSAD--EYELRVTRQD-EILFNSRSLSFETKRSV 131  
 Db 87 VLSNL-KPGIAGTIVLSIEMPAOSLTPROQYKLYIGETIENALIPENENELKYDOKALSV 145  
 QY 132 FIOTDKALYKROEVKRIYVLPSPKPKYKSLNLIKDKPKNLIQOMLSQOSDLGIVSK 191  
 Db 146 FIOTRAIRPAPSLVRYRAIVKSDLKPYGNATIKIPDSRNLISOTIGTILDRGVSG 205  
 QY 192 TQOLSHPIEDMSIQO---VNDQTYQSFQVSYVLPKFEVTLQTPYCSMSKHLN 247  
 Db 206 EYOLAEETLLGDMFLEVESNGVODKS---SFTVDYVLPKFEVNLKTSSTFTIND-DLS 261  
 QY 248 GTTAKYTYGKPYKGVDTLTF-LPLSFV-----GKKKN-----ITKTFKIN--GS 289  
 Db 262 VAVDAKTYTGKVAAGAKVSLPHEHRIAMPITIIDENGVKKEELMERVYKLNROCE 321  
 QY 290 ANFSFDEEMKNNVMSNGLSEYLDLSPGPEILTTVESYTGISRNVTWVFFKOHY 349  
 Db 322 AAVVFSNDELK-----HKLHMGWGGSIIRIVASVTEDEITELERNATHQISFREV 373  
 QY 350 IIEFDYTVLKPSLNFTAVKTRADN--QLTLEERNVYITVQNTYEWGSGNS 407  
 Db 374 KLDVEKOGDTPFKPGLTYNVVALKOMDDTPVKATLPKR---VOVSTFYNYF--YNHDS 427  
 QY 408 GNQKMAVOKINYVTPQSGTFKIBPILDESELQAKAFGLSGSKSMVHSLFK----- 461  
 Db 428 SLQEKETKIVEVDAGISVLTLOPINCTSARIEH-YDIGKDMFTATPYSSLYEA 486  
 QY 462 --SPSKTYIQLKTRDEN-IKVSPPELVVSGNKKLELSYMWVSRQOLV-----AVGKON 513  
 Db 487 AVSPKTSFLQLADNMGAVDVGKSLSPKATQPTLTIYQVMSNSNIIVSQOMTVNSH 546  
 QY 514 STMESLTPENSWTPKACVIVY-YIEDDEIISDVLPQVLPKNIKIKLYMSKVAEBS 572  
 Db 547 AAT-SFPATANMAPKSRILIVAIIESSQVLVADLDFKEGIFQNOVALISIKQAVEPQ 605  
 QY 573 KVSLSVTPQDSIYIVAVDKSVNLMNSNDITMENVHELEIVNTGYLLMF----- 626  
 Db 606 NKKEKVT-SDKNSPGLVAVDVSLLTKGNDITREKVRKODLENYDNNVGGFGGPRFW 664  
 QY 627 -----MNSFAVQECGLWLTDA-----NLT 647  
 Db 665 EALDRKKSIMRPMWGISGSDAQSIFSNAGLVLTALLYRREBQREMSERLANTPGILT 724  
 QY 648 KDYIDGVNDAVEYERFMEENEGHIVDHDLSLGS--PHVRKHPETWILDTNMGR 704  
 Db 725 VMMMGAPGMAEA--FAAPPMG-----GSSPPPTVRKFPPTHWIMSDIN-STS 771  
 QY 705 IYQEEVVPDSITSVNAAGFVISEDGLTTPPELOAFQPPFIFLPPSVIRGEF 764  
 Db 772 GEVEHEIEPDTITSVASTFAINEBNGVAPPTSKLVFPFPIQLNPAVVRGEK 831  
 QY 765 ALEITIFNLKATEVKKII--EKSDKPIILMTSRI--NATGHOQTLVSEDCAT-- 817  
 Db 832 ALLVVENYMEKEQDVTYLLKDKSGYLLKKDGVVARDEVQONRIVASAGGTSK 891  
 QY 818 -VLPPIRPHLGEIPITVTALSPTASDAVOTMLVYAEIGESYQSILDLTDNRLOST 876  
 Db 892 AVYFPFIVPSIGEIPHISAIASOGGDAVEMMLRVDPQYKVDRLNPIVDLNNSSDRS 951  
 QY 877 LKTISSFPNVTVTSERVOITAGVULGPSINGLASLIRMYGCGEOMINIPARTIYL 936  
 Db 952 -KNLEIIMVNDVVDGQKARLVLTGDMGPVNLNAHKLQVMPYGCGBOMMLLVPIIIV 1010  
 QY 937 DYLTKKQOLTMLKEKALSPMGQYQRELLYQREDSFAFGVYDPSGTSMTLSAFVLCF 996  
 Db 1011 KILRATNRESOLETAKIFIEGQIQRRELTLYKRADNSFSAFGSDKAGSTWLTAFVRSF 1070  
 QY 997 LEADPYIDIDQVNLHRTYTLKGHQKNSGEFMDPGVHISELQGG-NKSPVTLTAVIYTS 1055  
 Db 1071 HHAQVAFVDPNVISRAVAFILNSQOMESGAFARGBGVHKKDMQGGAGODGVALTAFVLIS 1130

QY 1056 LGYRKQPNIDVQSEIHFLSESRGISDNYTTLATLYALSSVSGSPRAKALNMLTWR- 1114  
 Db 1131 IL-----ENGENGAVATYLELKHIDEVSGNAVTVAAVVALQAKSKQAGAFELKHK 1185  
 QY 1115 -AEQEGKQFVWSE--SKLSDS---WQPSLDIEVAVALLSHLQFQTSSEGIPIRM 1167  
 Db 1166 IVEKSGDYKFAAQKVEKLEKESRAYMFOARPVIDETTSYAVLSYLAQNTSESISIRW 1245  
 QY 1168 LSRQNSLGGFASQDPTVALKALSEFALMTERNTIQUVY-TPSPSSP-----PLAY 1220  
 Db 1246 LVSQNELGFTSTQDVTMALQALSSYAAVYSDHNSQVTLNKGTHSPDINIRNAIV 1305  
 QY 1221 VQPM-----AVNISANGFPAICQLNVVNVNVAAGSSRRRSRSTIONEAPDLVAVKEN 1273  
 Db 1306 LQSYLSSLNDASVINANGTVFVQLSYST-----YRSLNDAPFCSEIET 1356  
 QY 1274 KODLNHVDLNVCTSPSGRSGMALMEVNLGSEVPSSEALSSETVK-----KVEYDHG 1328  
 Db 1357 RAG-NRLQDLCCVYTRGKSNMALAIDALSGRFPAEQVHTLTSIEDLQVMEKODT 1415  
 QY 1329 KLNLYDSVNETQFCVNIPAVRNKRVSNTQASVSYDYEPKQAVASVNSVYKLSGCD 1388  
 Db 1416 KKNVYFNLGGRPVCLSLYSQVTVQVADOKPANFRLVDYDPBOLKMTYAKQOTRSLQ 1475  
 QY 1389 LCSDVQGRP 1398  
 Db 1476 KCG-EDCW 1483  
 RESULT 6  
 08T398 PRELIMINARY; PRT; 1809 AA.  
 ID 08T398;  
 AC 08T398;  
 DT 01-JUN-2002 (TREMBlrel. 21, Created)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE Alpha-2-macroglobulin homologue.  
 GN A2M.  
 OS Clona intestinalis.  
 OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;  
 OC Phlebobranchia; Cloniidae; Clona.  
 OX NCBI\_TaxID=7719;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TTSUB-Hepatopancreas;  
 RA Hammond J.A., Nakao M., Yano T., Kemp G.D., Smith V.J.;  
 RT "Complete cDNA of an alpha2 macroglobulin homologue with a catalytic  
 histidine residue from Clona intestinalis."  
 RL Submitted (FEB-2002) to the EMBL/Genbank/DBJ databases.  
 DR EMBL, AJ431688; CAD24311.1; --  
 DR InterPro; IPR002890; A2M\_N.  
 DR InterPro; IPR001064; Crystallin.  
 DR InterPro; IPR001599; Macrogloblna2.  
 DR InterPro; IPR000884; TSP1.  
 DR Pfam; PF00207; A2M; 1.  
 DR Pfam; PF01835; A2M\_N; 1.  
 DR SMART; SM00209; TSP1; 1.  
 DR PROSITE; PS00477; ALPHA\_2\_MACROGLOBULIN; 1.  
 DR PROSITE; PS00225; CRYSTALLIN\_BETAGAMMA; 1.  
 DR PROSITE; PS50092; TSP1; 1.  
 SQ SEQUENCE 1809 AA; 199172 MW; E300DC643946E454 CRC64;  
 Query Match 24.2%; Score 1779.5; DB 5; Length 1809;  
 Best Local Similarity 30.0%; Pred. No. 2.3e-99;  
 Matches 453; Conservative 293; Mismatches 586; Indels 179; Gaps 41;  
 QY 13 LCVCTAALAAVAPGRFVLTARGIIRPGGNVITIGVLEHCSQVTVKALELKTASNLVVS 72  
 Db 16 LIATLSLATASNV-YNLYFPKHIRPGFNISFTALIDN-ENVQIHTAFPSMDSFHD 73  
 QY 73 VLEAGVEKSGFKTLTLPSPILN-SADEIYELVTRG---TODEILFNSNSTRLSFETKR 128

74 STDS--VNSGSSSRISNMGLPIHYSGSHGELNITGTDLVTGAQLFENSTDFQFQAKS 130  
 QY 129 ISVEIQTCDKALYKKEOEKFRIVTLFSDPKFKYKTSNLILDKPKSNLIQOMLSQOOSDLGV 188  
 Db 131 ISILIOTDKALYQGHVTKFRALAKLPDLKQONISITFDPRPGNVMLEPEVPLNHGV 190  
 QY 189 ISKTFQJSSHPILQDWSIQOVNDQTYQSFQVSEYVLPKREVTLQTPLYCSNMSKHLNG 248  
 Db 191 AGGQFSLTKDVAQMMKVEFMABGFKESLSVEKRYKLPKFKVVKAPSYIHPQSTGLTI 250  
 QY 249 TITKATYKGVKGDVLTFL-----PL--SFMKC-----KNITKTF-KINGSNFS 293  
 Db 251 KJDKAKTYFKGVQGGGLEEVGVGYQVPYHGFGRFAPRPTQKKITRRYPNFGTVELL 310  
 QY 294 FNDEMKVNDSSNGLSSEYDLSSPGVEILTTVTESVTGISRVNSTVNFPEKQDIIEF 353  
 Db 311 ITNEINEEL-CMGAGASESI-----ITVGSYTEALTRAFPDQRIAKTINVAVET 362  
 QY 354 FDYVTVLKPSLNFATYKVRADGNQTLTEER-RNNVITVQGN----- 397  
 Db 363 LVKELTITKPLGLKAYAIQITEVDGKPLPEDRLANNLLNIEYRPRGEPEGTNTTVST 422  
 QY 398 -YTEYVSGSNGNGKMEAVOKINTVPOSGFKIEPILBDS-SELQKAFGLSKSMA 455  
 Db 423 WYAYRMBETRV-----FVIPSGLVKTIDAPSDTFTSINFRPY--TNATMS 467  
 QY 456 VHSJLTK-----SPSKTYIQLKTRDENIKVSGPELVVSGNKLKELSVYVSGQLVA 508  
 Db 468 QRMALQMTARADSPNSYLOITTEBNSVPGNMATYIRTEAVESETLILIRGEILS 527  
 QY 509 VGR-----QNSTMESLTPENSWTPKACVIVYIIEDEGIIISDLKIPQVLFKKIK 560  
 Db 528 ERKFTQLSGVENSHLFEFVEYDMIPGVQVLASVRDGEIVADYIKLVTALLENQVS 587  
 QY 561 LYNKVKABEBEKYSIRISVTQPSYIGVAVADKSVNLMAANSNDITMENVHELELVNTG 620  
 Db 588 ITSSSTNIDAGEDVSIRVQTSSSGAYVGARIDSVLLKSGNVOSERIYTDLNKYSVT 647  
 QY 621 YLLCMF-----MNSFVFOECGLMVLTDANLTQDYIDGYD-----NAE 659  
 Db 648 QELNHRMRRMWWWPTPSGASDASDVFRKAGILVFTDLVYKPEASIIYPPPIAFSLNGG 707  
 QY 660 YAEFMEENEGHIVIDH--DFSLGSSP-HVRKHPETWIMLDTNMGRYIQOEFEVTVPDS 716  
 Db 708 FAER-----NIATAVNDTSTPATPRTRTLFPEETLMDEQJSGDAGSATFMTAPDT 760  
 QY 717 ITSVAWAGFVISEDGLGTLTTPPELQAFQFPFIPLPVYINGEPALEITTFNLIKD 776  
 Db 761 ITSMIFSASFVSDHGLGVSEQH-KVTVFRNFPTLMLPVAVINGELIIVQAIYFNLYST 819  
 QY 777 ATEVKVIEKDKPDILMTSSEINATGHOOTLLVPSDEGATVLPPIRPTHGELPIVTVA 836  
 Db 820 EVDVAVLTITESNKVLLRPGNNSAAGFSRRIITPASSSVKPIRMGTLGELPIMTVA 879  
 QY 837 LSPASDAVNTOMILVKAEGIEKYSOSIILDLTNRLOSTLTKTSFSPPTVTGSEBVQ 896  
 Db 880 ISEIASDALTRKVFQPEGITQCTSGSVLPQRMDASAPDVESLNIOIPAGIVGSEKVK 939  
 QY 897 ITAIGDVLGSPINGLASLIMRPGCGEONMNPANIYILDYTKKQOLDNLKEKLSF 956  
 Db 940 LLVVGDLIGSTMMNLGSLTRTSCCGEONMIGFAPDVFTLLYHSACKLDAATRAKAFKH 999  
 QY 957 MRQGYORELLYORSDGSFSAFGNVDPSGTSLASFVLRCPF---EADPYIDIDONVLRHT 1013  
 Db 1000 FQTSISELNKYNKHRDGSFASGEGASGTWLTLPAAFCFPAELRPTL-VASAVTDQA 1058  
 QY 1014 YTWLKHQKNGSEFMDPGRVHISELOGNKSPTVLTAYIVSLGYKRYQNPIDVQSEIH 1073  
 Db 1059 LTFILNQNTGTGTRERPGVSHKMOGQVDSPIMTAVLTLTKETNYAVANRAVQDAAE 1118  
 QY 1074 ----FLESSEFSGISDN-YTLALITYALSVSFPKAEKALMMLTWRAQBQSGMOFM--V 1125  
 Db 1119 NARIYENHILT-SISDNKYALAIYVALHVAGSSRAEALALBALATVQGGFKFMDNS 1177

QY 1126 SSESKLSDSWQ-----PRSLDIEVAAVALLSHPLQFQTESEGIPIMKMLSRQNSLGGFAS 1180  
 Db 1178 ESPDSYSRRRPPYVNPPTNDIEMSAVALLTYRRDLNAGIIPVMKWLASKRSSLGGYSG 1237  
 QY 1181 TQDTYALKALSLSFALMANTERTNIQVTVYTPGSSPEPL-----AVQPM--- 1224  
 Db 1238 TQDTYALIAQLSTYAGLVLVNTQNLQISASHSDPPTASYNINRENSIVNSVAVAVDG 1297  
 QY 1225 ANNISANGFPAICOLNVVNVKASGSSRRRSIQNQ-----EAFD-----LDVAVENKO 1275  
 Db 1298 TVQYATAGVAVAAQISVCNT-----PQPIEIEFQCTNVTYVSTALKAK- 1344  
 QY 1276 DLNHVDLNVCTSP-SGPRGSGMALMEVNLGFMVP-SEASISLSTYKVVYDHGKMLTY 1333  
 Db 1345 -----VNMCCSLRPGDMNATGMFLMEVNLPSGYTVINIDERTNPSAKLVEIDNGVNVY 1398  
 QY 1334 LDSVNE-TQPCVNIIPAVRNFKVSNTQDASYSIVDYIEPRRQAVRSNSEVKLSSCDLCS- 1391  
 Db 1399 YDEIAPGRSVCAIDIELNLGNVGSKARKVAADYVQPKERVVALYQVDEAPVVCDSCT 1458  
 QY 1392 -DVQGRPCED 1401  
 Db 1459 EDIYAVCSVCA 1469

RESULT 7  
 081PH5  
 ID 081PH5 PRELIMINARY; PRT; 1388 AA.  
 AC 081PH5;  
 DT 01-MAR-2003 (TREMblrel. 23, Created)  
 DT 01-MAR-2003 (TREMblrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)  
 DE CG7052-PE.  
 GN TEPIT.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NX NCBI\_TaxId=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Mortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.H., Blazer R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,  
 RA April J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballwey R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Botkova D., Botchan M.R., Bouck J., Brokstein P., Brotier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Palacios B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster G., Gabriellian A.E., Garg N.S., Gelbart W.M., Glaeser K.,  
 RA Glodde A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K., Howland T.J., Wei M.H., Ibegwam C.,  
 RA Jellali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasbo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Matei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy J., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
 RA Palazzolo M., Peltman G.S., Pan S., Pollard D., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Slier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weisenbach J.,  
 RA Williams S.M., Woodgett, Morley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
 RA Zeng R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195 (2000).  
 [12]  
 RP SEQUENCE FROM N.A.  
 RA Celinker S.E., Adams M.D., Krommiller B., Man K.H., Holt R.A.,  
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
 RA Banton J., An H., Baldwin D., Banton J., Beeson K.Y., Buesam D.A.,  
 RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
 RA Dodson K., Dorsett V., Doup L.E., Doyle C., Drenek D., Farfan D.,  
 RA Ferrieres S., Frise E., Galle R.F., Garg N.S., George R.A.,  
 RA Gonzalez M., Houch J., Hoskins R.A., Hostin D., Howland T.J.,  
 RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshirei A.,  
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nuno J.,  
 RA Pacleb J., Paragias V., Park S., Patel S., Pfeiffer B.,  
 RA Phoonanavong S., Pitterman G.S., Puri V., Richards S., Scheeler F.,  
 RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,  
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.,  
 RT "Sequencing of *Drosophila melanogaster* genome.";  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 [13]  
 RP SEQUENCE FROM N.A.  
 RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,  
 RA Tupy J.L., Bergman C., Bertram B., Carlson J.W., Celinker S.E.,  
 RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,  
 RA Krommiller B., Marshall B., Millburn G., Richter J., Russo S.,  
 RA Seale S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,  
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.,  
 RT "Annotation of *Drosophila melanogaster* genome.";  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 [14]  
 RP SEQUENCE FROM N.A.  
 RA Adams M.D., Celinker S.E., Gibbs R.A., Rubin G.M., Venter J.C.,  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 [15]  
 RP SEQUENCE FROM N.A.  
 RA Flybase;  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AB003618; AAN10639.1; -;  
 SQ SEQUENCE 1388 AA; 154276 MW; FE57D37D7DDCE1E CRC64;  
 Query Match 23.2%; Score 1703; DB 5; Length 1388;  
 Best Local Similarity 31.2%; Pred. No. 7e-95;  
 Matches 458; Conservative 288; Mismatches 539; Indels 182; Gaps 50;  
 QY 19 ALAAGPFLVTAAGIIRPG--NVTIGVELHCPQVYKAEILKTASULTVSVLEA 76  
 DB 14 ALLVATGIVSVVGPOTLRSNSKYNVAVSRADG-PSQ-IKVSINGPSYMET----- 64  
 QY 77 EGVFEKSPKTLTLPGLPLNSADEIYEL-----RVTRGTODELIFNSRLSPETKR 128  
 DB 65 -----KQIELP--PKSTQWVEFVKATATGNNTLSAEGSVGFNKSITKLVADKK 113  
 QY 129 ISVFTQDKALYKPKQEVKRIIVTLFSDPKPYKTS--LNLILKDPKSNLIQOMLSQOSDL 186  
 DB 114 PSVFOVTDKATYKPADLVQFRILFDENRPAKIEKPIVILIIIDGQNRKIKQLSDVPLTK 173  
 QY 187 GYISKTFOQLSHPIIDGMSIQOVN-DQYTYQSFOVSEYVLPKFEVTLTQPLCYSMNSKH 245  
 DB 174 GVPSSELOQISEQPVIGTWKISVSDGDNREKTSFEVDKYVLPKFEIYIVTTPRAVVIADKY 233  
 QY 246 LAGTITAKTYGKPVKGVDTLTF-----FLPLSFVKGKGNITKPKFKINGSANFSFNDRE 298  
 DB 234 IKATIRAKTYGKPVKGVKATVSMERSYGFGLDNLNANGKOE--KITIDVCKGHVEFD--- 288  
 QY 299 MKNVDDSSNGLSBYLDLSSPGVEILITVTESVTGISRNVSINVPEKQDYLIEFDYTT 358

DB 289 --IIHMAORG--QYLP-----PIKLFAVVTBELTGKKNKONATATVVLVHQORYSIEBERPE 339  
 QY 359 VLKPSLNFATVATKVRADGNOLTLERNNNVITVQRYNTEYWSGNSGKQKEAVOKI 418  
 DB 340 HFEANKSITIVVAVKANDGSPVT--NSAKVYKIGF-DKSYVFHPSPK-----TRI 388  
 QY 419 NYTVPOG-----GTFKIEPPILEDSELOLKAYFLGSKSSMAVHSLFK-SP-SKTYIQDKT 472  
 DB 389 NFEAPVNGIATFVRLPD-SDSRYYRIAFASPOSENTISISIKFEPTPMRSREPLKIQV 447  
 QY 473 RDNENKVGSPPELVSGKRLKELSYMVSNGOL-----VANGKONSTFSLTPENSW 525  
 DB 448 NTKRRLERQVSPDVASIEDLPFYFYITVARGANVILSDYVDPPOQKITYV-KFTPTFSM 506  
 QY 526 TPKACVITYYIEDGEIISDVLLKIPLYVFNKTKLVMSKYKAPSEKVSIRISVTOGDS 585  
 DB 507 VPKATIVYVAVVN-DLQFEKRTIDFEKFSNSIDVS-APINAKPSEVKLRIRK-TDADS 563  
 QY 586 IVGIYAVADKSVNLMAASNDITMENVHELRYNTGYLGMFNGFAVE--QECGLMVLTD 643  
 DB 564 FVGLIGVDQSVLLKSGNDLSQDDI FNSLNTYQTS---TPMNGYGRYPGQTSGLVLTIN 620  
 QY 644 ANLTQDYIDGV-YDNAEYAEKPEMBNESHVIDHDFSGSSPH--VAKHPEETWILDT 699  
 DB 621 ANPYNTDSISYD-----EVDAISITSTKIELVRYTNAEYVMMWTTTS 663  
 QY 700 NMGYRIVQEFVY--VPDSITSWATGPFVISEDGLGTLTPVLEQARQPFIFLNTLPS 757  
 DB 664 DNGSVGESEFVTKKIPITLTSWVATGSLPFTSGIALTKPKSKIRVQPPFVSTNLPS 723  
 QY 758 VIRGESEFALTEITFNYLKDQATEVKKILKSD-KEDILMTSE-----INATGHOQTLVLP 811  
 DB 724 VKRGVIAIPVIFNYLKDITADAVVMNSQOEVEFEATNEVLEKALIDEVRVRVRIIP 783  
 QY 812 SEDGATVLPPIRPHLGIPIPTVTLASPTASDAVTOMILVRAEGIEKSYSGSILDLTDN 871  
 DB 784 ANSGKSVFMRPKVKGFTTKITATSLADDAIHQKLKVEPEGVTLFENRAVFINLKDQ 843  
 QY 872 RLQSTLKLTSFPFNTYTGSEVQITAGIVLGPISINGLASLIRMPGCGEONNINAP 931  
 DB 844 PEMS--QSLDADIPNEVPOSEFIEFSVAGLGLPTQLNLDLVMPGCGEONNINAVP 901  
 QY 932 NIYIIDLVTKKKQOLTDNLKEKALSFMRQGYORELLYQREDSFSFAGVYDSDGFWLSAF 991  
 DB 902 NILVITKYLEVGRKLPKPSKAKKFLTELGYRELTYKHHDDSYSAFGSDASGFWLTAY 961  
 QY 992 VLRGFLADPYIIDQVYHARTYTWLKGHQKSGNEFMDPGFVHISELOGKNSPYTLTAY 1051  
 DB 962 VMRSFHQGTVDIDPKVITAGLDLFLVSKQESGEPFVGKL----FDNANQNPLALTSF 1017  
 QY 1052 IVTSLGTRKQPNID--VQESIHFLBSEFSGISDNVTALLITYALSSVSPKAKELN 1109  
 DB 1018 VLAFFENHELDIPKQSAIKKAVYVAEADK-TDDQSLIAVAVALQALAHGPOSEKIVA 1076  
 QY 1110 MLTWRAEOGQMOQFVSESKLSDS-----WOPSLDIEVAAYALLSHPLQFQSE 1160  
 DB 1077 KLESAVAKENBRMW-----STATESIGEDGVHFKKPRNSDVELITSYVLLLEKDPREK 1132  
 QY 1161 GIPWRWLSRQNSLIGFASQTDTTVALKALSEFA--ALMNTERTINIQVTYVGPB----- 1213  
 DB 1133 ALPIIKWILISQNSNGSGSTQDVIYGLQALTTFAYKSGSGTMDIFBSAAGESKNTIK 1192  
 QY 1214 -SPSPILAIVO-----PMANVISANGGPAICQNLVYVYNYKASGSSRRRSIQOQAFD 1265  
 DB 1193 VNPENSLVLQTHDLPKSTRKVDFTAKGSAWVQLSYRYNLAEK-----EKKPSFK 1243  
 QY 1266 LDVAIVENKDDLNHVDLNVCTSF-----SGPGRSGMALMEVNLISGFMVSESA---LSL 1316  
 DB 1244 VTPYVQDIPNQLLVD--VCAEYVPLBEDADKDSNMAVMIALPBGVVGSTISIGKIQ 1301  
 QY 1317 SETVAKVEYDH--GKLNLYDSVNETQF--CVNIPAVNFKVSNTOQDASVSIYDYPERRO 1373  
 DB 1302 VDRVAKVETKNSDSITVVVYFDLSITPGDVRCPLPESKKAHVAKQKPAVASLYDYDTBRK 1361



QY 473 RDNKVSPPFLVYSGNKLKELSYMVVSRGOL-----VAVGKONSTMSLTPENSW 525  
 Db 448 NTKKRCLEQVSPFDVVSIEDLPFYVYTVARGNVILLSDVVDVDPDQKTYTV-KFTPTESM 506  
 QY 526 TPKACVYVYIEDDEIISDVAKIPQVLFKMKIKLYMKAASEKVSIAISTQPS 585  
 Db 507 VPKATVYVYVNN-DLOPEEKTIPEKEFNSIDVS-APTNAKSEEVKLIK-TDADS 563  
 QY 586 IVGIYAVDKSVLNMNASNDITMENVHELELYNTGYGLMFPNNSPAVF--QEGCLMVLTD 643  
 Db 564 FVGLIGVQSVILLKSGNDLSGDDIFNSLNTYQTS---TPMANGYGRYVQGSGLVTLTN 620  
 QY 644 ANLTKOYIDGVYDVAERF-MEENEGHIYDHDSTL-----GSSPHVKHPPETWTL 697  
 Db 621 ANPY-----NTEPFIASFIAAQAIAQMPGTSIAHPNOAQAIRKEFPENWIF 672  
 QY 658 DT-MNGYRIYQEFYV--VPDSITSWATGFIYSELGIGLITTVELQAFPPFIPLNL 754  
 Db 673 NAEVNG---EEEFITLKIPDTITSMVYVGFSLNPTSGIALTKNSPKIRVFPFVSTNL 729  
 QY 755 PYSVIRGEFEALITFNYLKDTEVKVYIEKSD-KFDILMTSS-----INATGHQOTL 808  
 Db 730 PYSVIRGEVIALPVYIFNYLKDTEVDVMDSDDEYETETATNEXLEKALDEVRKRV 789  
 QY 809 LVPSDGAIVLPPIRPHLGEIPITVYALSPTASDAVTOMILVKAIEKYSQSILDL 868  
 Db 790 TTPANSKGSVSEFMIPKRVYFTLLKITATSLAGAHIOKLVBEQVGLTFENRAVFINL 849  
 QY 869 TDNRQSTLKTLSFEPRTYVGSERVQITAGDVLGSPINGLASIRPPYCGGQNMN 928  
 Db 850 KDQPMVS---QSLDADIPEVVPQSEFIEBSVVGDLGPTLQIMLVNRPYCGGQNMN 907  
 QY 929 FAPNIIYIDYLLKKQOLTDNLKEKALSFMRQGYORELLYOREDGSFSAQNVDPSPSTWL 988  
 Db 908 FVFNILVAKYLEVTKRKLPSVSKARKFLEIGYQELTYKHDGSAFCKSDASGSTM 967  
 QY 969 SAFVLRCLEADPYIDIDQVYHARTYMLKHKQKNGEFPDGRVISHLEQGNKSPVTL 1048  
 Db 968 TAYVRSFQAGTYIDIDPKVITAGDPLVSKQKSGEPVEVGL---FDNANQPLAL 1023  
 QY 1049 TAYVTSILGRKTYPNID--VQESIHFLESFSGISDNTYLLITVLSVSGPKAKE 1106  
 Db 1024 TSPVLAFFENHLLPKYQSAIKKAVRYAEADK-TDQVSLATAVAVALQAKHPOSER 1082  
 QY 1107 ALNMLTWAEQGGQMFVSSSEKLSDS-----NQPSLDEVAVALLSHFLOQ 1157  
 Db 1083 VIAKLESVARKENDRMM---SKATESTGEDGRVFMKPRSDVEITSYVLLALEKDP 1138  
 QY 1158 TSEGIPIRMWLSRQNSLGGFASDTDTVALKALSEFA--ALMNTERTNIQVTVGSPS-- 1213  
 Db 1139 AEKALPIIKMLISQNSNGSGFSSDTDYIGLOALTKFAVKGSSGTMIDIESSAGEBN 1198  
 QY 1214 ----SPSPVAVQ-----PMANISANGFGPAICQILNVYVNAKSGSSRRRSIQOE 1262  
 Db 1199 TIKVNPENSILVQTHDLPEKSTRKVDFTAKGTGSAMVQSYRYNLAEK-----EKKP 1249  
 QY 1263 AFDDLVAKENKDLNHYDLNCTSF-----SGGRSGMAMENVNLISGFWPBEA--- 1313  
 Db 1250 SFKVPFVKDPFNQLLIYD--VCAAYVPLEADAKDKKDSMAVMEIATLBPFGDSTSLGK 1307  
 QY 1314 ISLSTVKKVEYDH--GKLTLYLDSVNETOF--CVNIPAVRNKPVNTODASVYIDYEP 1370  
 Db 1308 IQAVNRVAVKRTKNSDSTVYVYFDSLTPDVACLPLEASKAAVAVKOKRPAVSLDYIDT 1367  
 QY 1371 RRQAVRSNSSEVKSJSCDLSV---QGR 1397  
 Db 1368 BRKATEYY--QVSSSLCIDICGADGCEGCK 1395

RESULT 9  
 Q81PH4 PRELIMINARY; FRT; 1408 AA.  
 AC Q81PH4;

DT 01-MAR-2003 (Tremblrel. 23, Created)  
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)  
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
 DE CG7052-PD.  
 GN TEPIT.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydriidae; Drosophilidae; Drosophila.  
 NC NCHI\_Taxid=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.H., Blazey R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabov G.L.,  
 RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brostein P., Brotlier P.,  
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieres S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hoston D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,  
 RA Jatalil M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Kerchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasro P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusken D.R., Patel J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Splitter B., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svrtkars R., Tector C., Turner C., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodgeret, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 "The genome sequence of Drosophila melanogaster.";  
 Science 287:2185-2195 (2000).  
 [2]  
 RN SEQUENCE FROM N.A.  
 RP Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,  
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
 RA Barton J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,  
 RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
 RA Dodson K., Dorselt V., Dou P.L.E., Doyle C., Dresnek D., Farfan D.,  
 RA Ferrieres S., Friese B., Galie R.F., Garg N.S., George R.A.,  
 RA Gonzalez M., Houck J., Hoskins R.A., Hoston D., Howland T.J.,  
 RA Ibegwam C., Jatalil M., Kruse D., Li P., Mattei B., Moshrefi A.,  
 RA McIntosh T.C., Moy M., Murphy B., Murphy L., Nelson K.A., Nunoo J.,  
 RA Patel J.B., Paragas V., Park S., Patel S., Pfeiffer B., Pfeiffer F.,  
 RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,  
 RA Stapleton M., Strong R., Svrtkars R., Tector C., Tyler D.,  
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;  
 "Sequencing of Drosophila melanogaster genome.";  
 Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.  
 [3]  
 RP SEQUENCE FROM N.A.  
 RA Mitera S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,  
 RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,

RA Ciamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,  
 RA Krommiller B., Marshall B., Millburn G., Richter J., Russo S.,  
 RA Seamlie S.M.J., Smith E., Shu S., Smutnick F., Whitfield E.,  
 RA Ashburner M., Gilbert W.M., Rubin G.M., Mungall C.J., Lewis S.E.,  
 RA "Annotation of Drosophila melanogaster genome."  
 RT Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RL  
 RN  
 RP SEQUENCE FROM N.A.  
 RA Adams M.D., Celiker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN  
 RP SEQUENCE FROM N.A.  
 RA FlyBase;  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AE003618; AAN10640.1; -;  
 SQ SEQUENCE 1408 AA; 156629 MW; A60FC3A3943431F4 CRC64;  
  
 Query Match 23.0%; Score 1693; DB 5; Length 1408;  
 Best Local Similarity 31.2%; Pred. No. 2,96-94;  
 Matches 461; Conservative 293; Mismatches 541; Indels 182; Gaps 54;  
  
 QY 19 ALAARPGRIYVTPGIIIPGCG--NVTIGVELLHSCSQVTVKELIKTASNLTVSLA 76  
 DB 14 ALANATGIYSVVGPTLRNSKYNVSVAKADG-PSQ--IKVSLNGPSTNET----- 64  
 QY 77 EGVEKSGFKTLTLPSPILNSADEIYEL-----RYTGRTODEILFSNSTRLSFETKR 128  
 DB 65 -----KOELRP--PMSTQNVFEPVKLATGNVNLASGCVGFVKSTKLTAVADKK 113  
 QY 129 ISVFITQDKALKYKROEYKPIVTLFSDFKPYKTS--LNIILKPKSNLLOQMLSQSDL 186  
 DB 114 PSVFVQIDKATYKPADIVQFRIFLDENTRPALKEKPISVIILIGQNRIRIKQLSDVLT 173  
 QY 187 GVISIKTQLSHPLIDKMSIOOVN--DQTYVQSFQVSEYVLKPEVTLQEPVLYCSMNSKH 245  
 DB 174 GVFGELQLSBQVPLGTMKLSVSDGDRKTSFVVDKTVLPKEVLYDTPKAVIVADKV 233  
 QY 246 LNGITTAAYTKGKVKVGDVTLT-----FLPLFWGKKKNIITFKINGSANFSDDEE 298  
 DB 234 IKATIRAKTYTKGVKVKATYSMERISYCFDGLNANGKOR--KTIDVDGKHVEFD--- 288  
 QY 299 MKNVMSNGSLSEYLDSSPGPVEILTTVTESVTGISHNVSTNVPKQHDYIIEFDYTT 358  
 DB 289 --LIHMAORG--QYLP-----PIKLFVAVTEELTGKQNAATATVVLHQRYSIERYEPE 339  
 QY 359 VLKSLPLATVKTTRADGNQTLLEERNNVIVITQNTYEVSGSNGQKMEAVQKI 418  
 DB 340 HFEANKSFITYOVVAVKAVDGSPT--NSAKNVKIGP-DKSYSEFHEPSPK-----TRI 388  
 QY 419 NYTVAPQS-----GTFKIEPPILEDSESELQKAYFLGSKSMVAHSLFK--SP-SKTYIQLKT 472  
 DB 389 NFEAPVVENGIATFNVRLPD--SDSRYYRIIFASFGSENTISIKFEETPMSRREPIKQV 447  
 QY 473 RDNIKVSGPELVSGNKRILKELSYVNVSGQI-----VAVGKQNSTWFSLTPENSW 525  
 DB 448 NTKKPRLOEQVSFDVVSIEDLPYFYVYITVARGNVILSDYVVPDQKQYTV--KETPFPSM 506  
 QY 526 TPKKACVIVYIEDDGEIISDVLEKIPQVLPFKNKIKLWMSKAKARSEKVSRIIVTQDS 585  
 DB 507 VPKKTIYYVYVNN--DIQFEKRTIDFEKEFNSNIDVS--APINAKPSEEVKRIK--TDADS 563  
 QY 586 IGVIAVADKSVNLMAANDITMENVHLELYNTGYLGMFNSPAFV--OECGLMVLTLD 643  
 DB 564 FVGLIGVQSVLILKSGNDLSODDIFNSLNIYQIS---TPMANGKGRPGQTSGLVLTIN 620  
 QY 644 AN--LTKDYIDGVYDNAS--VAERFMEENGHIIVDHFSLGSSP-----HVRKH 690  
 DB 621 ANYPNTTFPVDDEPEIYA---FENNLDALPMP--AIANFPDGTNTVQVPEIRKNF 675  
 QY 691 PETVWLTDMNGYR--YQEFVY--VPSDITSWATGVVISDGLGTTTPVELQAFQ 747  
 DB 676 ADVIMW--QSIGRSGVEEFLTKIKIPDITISWVTVGSLNPTSGIALTKNPSKIRVQ 733

QY 748 FFIPLNLPYSVIRGEFEALETIFNLYLKDATEVKVILKSD--KFDILMTSS-----INA 801  
 DB 734 FFIYTLNLPYSVIRGEFEALETIFNLYLKDATEVKVILKSD--KFDILMTSS-----INA 801  
 QY 802 TGHQQLVLPSESGAVLVPRIPTHGELPIYVTAASPAADVNTQMILYKAGIEKYS 861  
 DB 794 VRRKRVTLTPANSKGVSPFIRKNGVFTTLTKTATATSAAGDAIHOKLKVEPGLVLFEN 853  
 QY 862 QSILDLTNRLQSTLTKTSFSEPPNTVTSSEKVOITAGVLPISINGLASLIRMPYGC 921  
 DB 854 RAIFINLQKQPMSS--QSLDADIPNEVPOSEIERIESVGDLLGPTLQNLNDLVNRPYGC 911  
 QY 922 GEONMINPAPNIYILDYLTKKQLTDLNLEKALSPKQGYORELYQREDEGSFSAFGND 961  
 DB 912 GEONMNVFNNIILVLYLEVTGRLKPSVESKARFLEIGYQRELYTKHDDGSYSARFGSD 971  
 QY 982 PGGSTWLSAFLVLCFLEADPYTIDQNLVARTYTLKHQKSGSGEPMRGRVHSELQGG 1041  
 DB 972 ASGSTWLTAVWMSFFQAGTYTIDIPKVTAGIDFLVSKQESGEPVEGKL---FDNA 1027  
 QY 1042 NKSPVTLAVYVTSLLGKRYQCPNID--VOESIHFLSESESRGSDNYTLATITVLSV 1099  
 DB 1028 NQNPPLALTSFVLLAFPENNELIRPKYSALIKAVRYVAEADK--TDQOYSLAIAVLAQL 1086  
 QY 1100 GSPKAEALNMLTWABQEGGQFWSSSEKSDS-----KOPRSIDIVAAVALL 1150  
 DB 1087 KHPQSEKVIKLESVARKENDRMW-----SKATESGDEGRVHMKPRSNDEITSVLL 1142  
 QY 1151 SHFLQPQTSRGIPIMWLSQNRSLGSPASTOOTVTAALKASBFA--ALMNTERTNIQVT 1208  
 DB 1143 ALLEKQPAEKALPIIKMLISQNRNSGFSSTQDTVGLQALTFKAYKTGSGSGTMDIEFS 1202  
 QY 1209 VTGPS-----SPSPPLAVQ-----PMANVISANGFGFALCOLNVVNVYKASGSSRRR 1255  
 DB 1203 SAGESKRTITVNBENSLVLOTHLPKSTRKYDPTAGTGSAMQLSRYVLAKE----- 1256  
 QY 1256 RSIQNEAFDLVAVKENDLNLHVDLNVCTSF-----SGPRSGMALMEVNLISGFNV 1309  
 DB 1257 ---EKPSFVYTPVYKDTPNQLIIV--VCAEYVPLEDADKDKDSNMAVMEIALPSGFVG 1311  
 QY 1310 PSEA---ISLSEVYKVEYH--GKNTLYLDSVNEQF--CVNI PAVNPFVSTQDASVS 1363  
 DB 1312 DSTSLGIQAVDVKRVEYNSDSTVAVFDSLTPGDVRLPLEASRAVAVALQKSPASVS 1371  
 QY 1364 IVDYEPBPQAVRSYNSSEVLSSCDLCSDV---QGR 1397  
 DB 1372 LYDIYTERKATERY--QVKSILCDICEGADCEGCK 1406  
  
 RESULT 10  
 Q9NFV7 PRELIMINARY; PRT; 1420 AA.  
 AC Q9NFV7, Q9YLX8;  
 DT 01-OCT-2000 (Tremblrel. 15, Created)  
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
 DE TEB2 protein precursor (CG7052 protein) (SH01829p).  
 GN TEB2 OR TEB2 OR CG7052 OR CG18589.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_Taxid=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Laqueux M., Petrodou E., Levashina E.A., Capovilla M., Hoffmann U.A.;  
 RT "Constitutive expression of a novel complement-like protein in Toll  
 and Jak gain-of-function mutants of Drosophila."  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkely;  
 RX MEDLINE=20196006; PubMed=10731132;





Db 564 FVGLGVDQSVLLKSGNDLSQDDI FNSLNTYQTS---TPMNGXGRYPQGTSGVLTTLN 620  
 Qy 644 AN-----LTKDYI-----DGVYDNAMEYARFME--NEGHIYDIDHFS 679  
 Db 621 ANYRYNGLPVMYSVFEBSRHPMTTRRYKRGIGDSGDRISFISQSLNDKLEI---L 677  
 Qy 680 LGSSPH---VAKHPPTWIMLDITMGKRIYQEFVLT--VPDSITSWATGFGVISEDGLG 734  
 Db 678 LKQPPQRTTIKEPPEWTF--ENVG---EEFPLTKIKIPDTIISWVATGSLNPTSGIA 732  
 Qy 735 LTTTPVELQAFQPFIFLNPYSYIRGSEFALTEITFNYLKDAIEVKYIEKSD-KPDIL 793  
 Db 733 LTKNPSKIRVQPFVSTNLPYSVKGGVIAIPVIFNYLDTKLTDADVMMNSQGEYEFT 792  
 Qy 794 MTSSE-----INAGHOOTLLVPSDEGATVLPFIRPTLGEIPITVATLSTASDAYTOM 848  
 Db 793 EATNEVELEKAIDEVRKRVITIPANSKGSFVPIRPNKVGTTIKITNTSLAADAHOK 852  
 Qy 849 ILVKAEGIEKSYQSILDLTDNRLQSTLKTLSFSPPNTVSGSEVQVQITAGDVGPSI 908  
 Db 853 LKVEPEGVTLPEHNAVFINKDQPEMS--QSLDADIPHEVVPQSEFIEFVSGDGLGPTL 910  
 Qy 909 NGLASLIRMPYGCCEQNMNINAPNITIIDYITKKQLTDLNKEKALSPMKGIORELLYQ 968  
 Db 911 QNLDNLVMPYGCCEQNMNINAPNITIIDYITKKQLTDLNKEKALSPMKGIORELLYQ 970  
 Qy 969 REDSFSAFNGYDSDGSTMLSAFVLRCPLEADPIDIDQNTVHTYTWMLKCHOKSNGEPW 1028  
 Db 971 HDDSSTAFGKSDSGSTMTLTAWRSHQGYTTDDPKVITGDLPLVSKQSESEFP 1030  
 Qy 1029 DPGFVHISELOGKNSPVTLLTAIVTSLGKYKYOPIID--VQBSIHFLSESPRGISDN 1086  
 Db 1031 EVGKL-----PDNQNPLALTSFVLAFEFENHEILPKYQSAIKKAVRYVAEADK-TDDQ 1085  
 Qy 1087 YTLALTYLALSSVSPRAKELMNLTWABEGGKQFVNSEKLSDS-----WQP 1137  
 Db 1086 YSLAIAVAVALQALAPQSEKVIKLESVAREKENDRMW---SKATESYGEDEGVFHKMP 1141  
 Qy 1138 RSLDIEVAAYVALSHFLOFOTSEGIPIRMWLSRONSLIGFASQDPTVALKALSEFA-- 1195  
 Db 1142 RSNDEVTSTYVLLLEKDPKPKALPIIKLISQBNNGSSQDVTYIGLQALTKRAYK 1201  
 Qy 1196 ALMNTERTNIQVTVTPS-----SPSPPLAVQ-----PMAVNISANGFPAICQLAN 1242  
 Db 1202 TGSSTGMDIEFSSAGSSKNTIKVNPENSLYLQCHDLPKSTRKYDPAKAGSAMVQLSY 1261  
 Qy 1243 VYNYKASGSSRRKRSIONQEAFLDVAKENKDLNHYDLNVCISF-----SGPGRSGM 1296  
 Db 1262 RYNLAER-----EKKSQFKVTPYKDPNQLLIVD--VCAEYVPLEBDADKXDSNM 1310  
 Qy 1297 ALMEVNLISGFMPVSEB---ISLSEYTKKVEYDH--GKMLNYLDSVNETQF--CVNIPAVR 1350  
 Db 1311 AMEIALPISGFGVSTSLGKIQAADVRYKVEYTKNSDSTVAVYFSLTPGQVRCPLPBAK 1370  
 Qy 1351 NFKVSNTOQDASIVDYEPERRQAVRSYNSVKSLSDDLCSDV---OGCR 1397  
 Db 1371 AHAVALKQKPSVSVLYDYDTERKATEY--QVKSLSLDIGGAGCGSGCK 1418  
 RESULT 11  
 Q9VLY9 PRELIMINARY; PRT, 1399 AA.  
 AC Q9VLY9  
 DT 01-MAY-2000 (TREMblrel. 13, Created)  
 DT 01-OCT-2002 (TREMblrel. 23, last sequence update)  
 DT 01-MAR-2003 (TREMblrel. 22, last annotation update)  
 DE CG7052 protein.  
 GN TEBII OR CG7052 OR CG18589.  
 OS Drosophila melanogaster (fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NCBI\_Taxid=7227;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkely;  
 RX MEDLINE:20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazer G.G., Champagne M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Baas A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Botkova D., Botchan M.R., Bouck J., Brocktein P., Brothier P.,  
 RA Burdick K.C., Busam D.A., Butler H., Cadiett E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Fertiera S., Fleischmann W.,  
 RA Foster C., Gabriellat A.E., Garc N.S., Gelbart W.M., Glasser K.,  
 RA Glisdek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ileguam C.,  
 RA Jatalai M., Kalush F., Kapten G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasbo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Murty D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Slater E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "the genome sequence of Drosophila melanogaster";  
 RL Science 287:2185-2195(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Celniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,  
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
 RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,  
 RA Carlson J.W., Center A., Champagne M., Davenport L.B., Dietz S.M.,  
 RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
 RA Fertiera S., Frise E., Galle R.F., Garc N.S., George R.A.,  
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,  
 RA Ileguam C., Jatalai M., Kruse D., Li P., Mattei B., Moshrefi A.,  
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,  
 RA Pacle J., Paragass V., Park S., Patel S., Pfeiffer B.,  
 RA Phoumenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,  
 RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,  
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;  
 RT "Sequencing of Drosophila melanogaster genome";  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Miera S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Prochorko S.E., Smith C.D.,  
 RA Tuhy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,  
 RA Ciamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,  
 RA Krommiller B., Marshall B., Millburn G., Richter J., Russo S.,  
 RA Seale S.M.J., Smith E., Shu S., Smutnick F., Whitfield E.,  
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mangall C.J., Lewis S.E.;  
 RT "Annotation of Drosophila melanogaster genome";  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter J.C.;

RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA Flybase;  
 RL Submitted (SEP-2002) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; AE003618; AAF52539.2; -.  
 DR HSBP; P01024; IC3D.  
 DR Flybase; FBgn0041182; Tepit.  
 DR InterPro; IPR002890; A2M.N.  
 DR InterPro; IPR001599; MacroglloblnA2.  
 DR InterPro; IPR002125; Serpin.  
 DR Pfam; PF00207; A2M; 1.  
 DR Pfam; PF01835; A2M.N; 1.  
 DR PROSITE; PS00477; ALPHA\_2\_MACROGLOBULIN; 1.  
 DR PROSITE; PS00284; SERPIN; 1.  
 SQ SEQUENCE 1399 AA; 155783 MW; 3D8B7DF91E39F59 CRC64;

Query Match 23.0%; Score 1688.5; DB 5; Length 1399;  
 Best Local Similarity 31.3%; Pred. No. 5.4e-94;  
 Matches 461; Conservative 285; Mismatches 544; Indels 183; Gaps 52;

19 ALAVAPGPRFLVTAAGIIRPGG--NVTIGVELLEHCPQVNTKAEILKTAASLTVSVLEA 76  
 14 ALAVATGTVSVGCGTLLSKNSKTNVAVSHKADG-PSQ--IKVSLNGPSYNET----- 64  
 77 EGVFKEGSEFKTLTLPSPILNLSADEIYEL-----RVGRTQDEILFESNSTRLSFETKR 128  
 65 -----KQELP--PMSTQVNEFEVPGKATGNVNLASBEGSVGFNGSTKATVADK 113  
 129 ISVFQITDKALYKPKQOEKVRIVTLPSDFKPKYTS--LNLIKDKPSNLIQGLMSQSD 186  
 114 PSVFQITDKATKPADLVQFRILFLDENTRPKIEKPIISVIIIDAQNRIRKIQSDVLT 173  
 187 GVISTFQLSHPIIGDMSIOYQVN--DQTYGSGQVSEVLPKPEVTLQTPICYSMSNGH 245  
 174 GVFSGELQISEQVLTGTXISVSDGDNRETSFEDKTVLPKFEVITVTPAAVVIADKV 233  
 246 LINGITAKTYGKPKVGDVTLT-----FLPLSEWKKKNTTKFKINGANSFENDEE 298  
 224 IKATIRAKTYGKPKGKATVEMERSYGFQDLNNGKQE--KIIDVGKQHVDFD--- 288  
 299 MNVMDSSNGLESEYLDSSPGVEILTVTTSVGTISRNVSNTVFEFKQHDYIIIEFDYTT 358  
 289 --IIMAGQG--QYLP-----PIKLFAYVTEELTGKQNAATVAVLHQGRYSIEPERE 339  
 359 VLKPSLNFATYKATRADGNOULTLEBRNNVITYTQKNYTYMGSGNSGNGKMAVQKI 418  
 340 HFEAKSFLYQVAVKNVDSPTV--NSAKNVKIGF-DKSYSYFHEBSPK-----TRI 388  
 419 NTVFPOS-----GTFKIEPILDESSELQKAYFLGSKSMVAHSLFK-SP-SKTYIOLKT 472  
 389 NFEAVNENGATITFNVRLPD--SDSKRYRIFASFDSSENTIGISKEFEPKMSREPLKQV 447  
 473 RDNIRKVGSPFELVSVGNKRLKELSYMVVSRQL-----VAVGKQNSTMSLTPENSM 525  
 448 NTKKRLPGQVSEFVDSIEDLPYFYTVIARGNVILSDVVDPDQKTYTV-KFPTPTSM 506  
 526 TPKACVIVYIIDDEIISDVAKIPQVLFPKAKIKLYSKAKAASEKSLRISTYQPS 585  
 507 VKKATIVYVVVNN--DLQFEKTIIFEKKFSNSIVS--APTAKSSEEVKRIK--TDDDS 563  
 586 IYGIYAVDKSVNLMAANSNDITMENVVELELNTGYLGMPNNSFAVF--QEGCLAMVLT 643  
 564 FVGLGVDOSVULLISGNDLSODDIFNSLNTYQTS---TPMNGIGRYFGQTSGLVTLTN 620  
 644 ANLTKDYIDGVYDNMEYAEERMEENEGHIVDHD--FSLGS-----SPHYRKHFPEW 694  
 621 AN-----YPYTERIRIFRPGIGFPRPLFNRYTVAGSLPPNVIPRQYKPEPEW 671  
 695 IM-LDTNMGKRIYQREVT--VPDSTSVNATGCVISELGLGTLTPPELQAFQPFPT 751  
 672 IENIFENVG---EEEFLLTKKIPDITTSVVVGFSLNPTSGIALTKNPKSKIVQFPFFVS 728

QY 752 LNLPSYVIRGEFALEITITFENVLKDAVEKVIIEKSD-KEDILMTSSE-----INATGHQ 805  
 Db 729 TNLPSYVIRGEVIAIPVYIFNYLDKTLDAVMDNSDQVEFTEATNEVLEKALIDEVARY 788  
 QY 806 QTLVPSDQATVLPFRPHILGEIPITVNLSTPAQVAVQMLVYKAEIGEKSYQSIL 865  
 Db 789 KRVTIPANSKGSVSEFMIRPKNVGFTTLKITVTSALAGALHQKIKVEEGVTLFENRAVF 848  
 QY 866 LDLDNRLOSTLTKLTSFSPFNTVTGSEERVOITAGVLSINGLASLIMPYGCGEON 925  
 Db 849 INLQDPEMS--QGLDADIPHEVYPOSEFIEFSVGDGLGFTLQNLNLVMPFGCGEON 906  
 QY 926 MINEPANIYIIDLVTKKKQLTDLNKEKALSEMRQGYORELLYQREDSFSAFGVDPGS 985  
 Db 907 MNVEPNILVLYKEVTRGKLPVSESKARKPELIGYQRELTYKHDGDSGSAFGSDASGS 966  
 QY 986 TWLSAFVIRCLLEADPIYIDQNVLHRTYTLKQHSKNGSEFMPPGKVIHSELQGNKSP 1045  
 Db 967 TWLTAAYVRSPHQAGTYTIDIPKVIYTAGLDFLVSKOKSGEFPFVGKX----FDNANQNP 1022  
 QY 1046 VTLTAIVTSLGIRKQOPNID--VQSIHPLSEFSGISDNTYTLALITYALSSVSPK 1103  
 Db 1023 LALTSFVLAAFFENHELIPTQSAIKKAVRYVABADK-TDDQYSLAIAVALQLAKHPQ 1081  
 QY 1104 AKKALNMLTWRAEAGEGQWQVWSSSEKLSDS-----WQPSLDIEVAAYALLSHFL 1154  
 Db 1082 SEKVIATKLESVARAKENDRMW---SKATESTGEDGRVFMHKPKPSNVEITISYVLLALE 1137  
 QY 1155 QFOTSEGIPIKRWLSRQNSLIGFASQTDTVALKALSEFA--ALMNTERTINQVTVGP 1212  
 Db 1138 KDPAEKALPIIKWILISQNSNGGFSSTODTYIGQALTKFAKYKTSGSGTWDIFFSSAGE 1197  
 QY 1213 S-----SPSLPAAVO-----PMANISANGGFPALCOLNVVYNYKASGSSRRRSIQ 1259  
 Db 1198 SKNTIKVNPENSLVQLTHDLPKSTRKVDFTKGTGSAVQVLSRYNLAEK-----E 1248  
 QY 1260 NQEAFLDVAVXKENDLNVHDLNVCTSF-----SGPGRSGMALMEVNLISGFMVSEEA 1313  
 Db 1249 KKSFKVTPYKQDPNQLIYD--VCAEYVLEADADKQKSNMAVMEALPLSGVIGDSTS 1306  
 QY 1314 ---ISLSETVKRYEYDH--GKLNLYLDSVNETQF--CVNIPAVRNFKVSNTOASVSYDV 1367  
 Db 1307 LGKIQADVRRVRETKNSDSSTVYVYFDSLTPFGDVRCLPLEASKAAVAKQCPASVSYDV 1366  
 QY 1368 YEPRRQARSYNSSEVYKLSGCDLCSGV---QCCR 1397  
 Db 1367 YDERKATEY--QVKSGLCDICGADCGEGCK 1397

RESULT 12  
 ID 001717 PRELIMINARY; PRT; 1507 AA.  
 AC 001717;  
 DT 01-JUL-1997 (TREMBLrel. 04, Created)  
 DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
 DE Alpha-2-macroglobulin.  
 OS Limulus sp.  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Merostomata; Xiphosura;  
 OC Limulidae; Limulus.  
 OC NCBI\_TaxID=51645;  
 RN  
 RP SEQUENCE FROM N.A.  
 MEDLINE=97175055; PubMed=9022715;  
 Iwaki D., Kawabata S., Miura Y., Kato A., Armstrong P.B.,  
 RA Ougley J.P., Nielsen K.L., Dolmer K., Sottrup-Jensen L., Iwanaga S.;  
 RT "Molecular cloning of limulus alpha-2-macroglobulin";  
 Eur. J. Biochem. 242:822-831(1996).  
 DR EMBL; D83196; BAA19844.1; -.  
 DR HSBP; P01024; IC3D.  
 DR InterPro; IPR002890; A2M.N.  
 DR InterPro; IPR002048; EF-hand.  
 DR InterPro; IPR001599; MacroglloblnA2.

DR Pfam; PF00207; A2M; 1.  
 DR Pfam; PF01835; A2M; 1.  
 DR PROSITE; PS00477; ALPHA 2 MACROGLOBULIN; 1.  
 DR PROSITE; PS00018; EF HAND; 1.  
 DR SEQUENCE 1507 AA; 168490 MW; 88F4013495F63CED CRC64;  
 Query Match 22.3%; Score 1638; DB 5; Length 1507;  
 Best Local Similarity 28.9%; Pred. No. 7.4e-91;  
 Matches 449; Conservative 270; Mismatches 559; Indels 278; Gaps 43;  
 28 FLVTAAGTIRRGQVNTIGVLELHCPQSVYTKAKLLTASNLVTS-----VLEAAG 78  
 29 FLTAAPSLTPGKSLIHLPLD-----IKNGFLRIGVQDDQDDVAAETEV 76  
 79 VFEK---GSFKTLTPS-----LPLMSADEIYELRVGTODELIFSNSTLSPEFRKIS 130  
 77 SFNDNPSSSIQILNIPSGEVEKPKLVANGSY-----SSPSNDPFPEKD--INNHOKLI 130  
 131 VFIOTDKALYKPKQVVKFRIVTLFSDFK--DYKTSNLILKDPKSNLIQOQLSQOQDLGV 188  
 131 VFVQTDKPLKPGQTVKVRILPTTPDLKLVKETIGSFQIENPDGIVLGYPMLSFAGI 190  
 189 ISKTFQLSHPIILDMSIQOVNDQTYQSFQSEVYLKPEVYLLQTPLYSNMNSKILNG 248  
 191 AQFELALPDEPTYGMMRIKNIIDTEIYENFEVKEVYLPKEVKITPESYLLTNADSITW 250  
 249 TITAKYTGKPVKGDVTLTFLPLSFMGKKNIITKTPK-----INGSANFSFND 297  
 251 KICQYITGQVPEGTFAVETNVKKNKEGVPIIHKEGLDGLDLYVASSA--LGNEQ 309  
 298 EMKQVMSNGLSEYLDLSSPGVEILTTVESVTGSRNVSTNVFQKHDIIEFPDYT 357  
 310 RL-----SYRAVMFEALVEKGTGIKMATDSDIYRTSNPLNIMYLEPT 352  
 358 T---VLRSLNFTATVKTADGNQOLTEERRNNVITVQNTVTEWSSGNSNQMEA 414  
 353 SGKYLKPLKPLPGYGLKVEKPDGTPAPGEQ-----LELCFADREKMRNRKMLEEKIRA 406  
 415 VQKINTVPOSGTKIEF--PILEDSELOLKAFL-----GSKSMAYVSL---FKS 462  
 407 CKE--FTSDENGIIKFTVPPOTPDITSRFKAKLQYGGKDGDKLNQPHSFLVSSWYS 464  
 463 PSKTYIQLTRDENIKVGSPEL---VVGSKRLKELSYVNVSGQLVAVGK----- 511  
 465 PSGHLOLEPITEIECGKPLTVKFKYTTGEEKQKQFVYQIMANFIYDGSFHEFLIS 524  
 512 -----ONSTMFSLTPEN--SW-----TPKACVIV 533  
 525 EDKSGLTDETYLPIDVTLALNPNEPEWENNVIYPHIGETSLTLIPSEFMNSAKILV 584  
 534 YIISDGEIISDVLIKIPQVLFKKNKILKYMSKVAEPSEKSLRISTVOPDSIYGIYAVD 593  
 585 FYVDEDETVADSTKITYKKLKRNVGLKFGEEKVLFGASTLQTLTAS--PISIGIAYD 643  
 594 KSVNLMAASNDITMENVVHEL-----ELVNTGY 622  
 644 KSVHILSSDKRITTEEBVYKLGHDYWPQKATSDYKCYCEDYKQTEGHEGSSFGFT 703  
 623 LGMFMNSFAVQEGGLWTLTANL-----TKDYIDGVYDVAEAEKFMENEGHIV--- 673  
 704 STYNLDITATPDEGLVVISDMELETRPCKSGFGEDEPCQPDVAFAPQANRIGGG 763  
 674 -DIHDFSLG-----SSP--HYRKHPEPTWILDTNMGRIYQREVVVPSITSWVTGF 725  
 764 GEAGGFGGGRIRKTKNKPVEIRITVFPETWLMELONIGATGELSKRIPIPTITWVSAI 823  
 726 VISEDGLGTTTVEVLAQPFPIPLNLPSYVIRGEFALEITIFNYLXDAEKKYIIE 785  
 824 CISETGAGVSEA--ATYKGFQPFVSTFLPSVIRGEKVPITVYVNYLSLECPITKLSLE 882  
 786 KSDKFDILMSSSELNATGHQOQLTVPSBEDCATVLPPIRPHLGIPIITVVALS----- 838  
 883 QSDKF-----EMQNDTNSYTS-CVCGSGKSDTTRMWMIKPRSLAGVNLTVGASLPEAIC 935

QY 839 -----PTASDAVYTMILVKAEGIEKSYSSOSILDITDNRLOSTLTKTISFSPPTVYG 891  
 DB 936 GNDQSYTVARDAATQQLVVEPEGFPEKEDTWSTFACPKQNGKFT--ATSDLLPELVED 994  
 QY 892 SERVQITATGVDVGSINGLASLIRMPYCGGEQNMNFAPNIYILYTRKKOULTNKE 951  
 DB 995 SARGVSYIGDLMGPAIKNIDHLVRLPTCGGEQNMVFPVNIYVLDYLTATGSIITISKE 1054  
 QY 952 KALSPKQGYORELLAQREDSFSAFGNVPDGSSTLSAFVLRCFLEADPYIDIDONVLA 1011  
 DB 1055 KALNNRKYARQNYRHPDGSYSAFGNDKQGNLFTLVFVRSFQAERFILNNKLN 1114  
 QY 1012 RIYTWLKGQKSGEFPWDPGRVHSELQCG---NKSPTVLTAVYVTSLL--GYRKYQPN 1065  
 DB 1115 ETENWILNQRNSNGCRKIKGLFNSALKGISNDPTPAFLFAYVILSLIAGYKQ---- 1170  
 QY 1066 IDVQESIHPLSEFSFGIS-----DNYTLALITYALSSVGSPPAKALNLTARAQ 1117  
 DB 1171 -----ETVLDGISCLEALSNPSTYSIALFAYATSLAGHPAKQYLALEERRAIT 1220  
 QY 1118 EGGMOFWVSESKLSDSQPRSLDIEAAYA---LISHFLQFQTSBGIPIMRWLSQRNS 1174  
 DB 1221 EGSKTFWKSPPS--GRYWGNSIGVELAGYAVLTLQHGASNLAKVTHIIRMLAQQNY 1278  
 QY 1175 LGGFASTODTTVAKALSEFALMINTERNIQVTVTPSPSPPLAVVQ----- 1222  
 DB 1279 RGGFSTQDTVIALQMSKFAITIIYDELDELVGEVSSGFEKKIMLTCONSILMQTFRLQ 1338  
 QY 1223 --PMAVNISANGGPAICQLANVYTNKAGSSRRRSRSTIONQAEFDLDVAVKEN--KQDLN 1278  
 DB 1339 TVSPVPDFEATGSGCLVQTSLRVNTP-----PPRGFLHEVTVKGLYBDCIN 1389  
 QY 1279 HVDLNVCTSPSGPGR--SGMALMEVNLISGFMPVSEAIIS-----LSRTVKVEYDHQKLN 1332  
 DB 1390 -AHIACTVKYDKGAGVSNMVALEMKNVSGWIPDEBSIKIIVDREELNIRYEDGQQLND 1448  
 QY 1333 YLDSVNETQCVNIPARNFKVSNTOASVSIYDYEPRAOAVRSYSEVKLSSCD 1388  
 DB 1449 YPSELTDQMLCFMFMEQDIEVQETKPAIRLYDYELQEVVTSYSID---ENCE 1501  
 RESULT 13  
 Q9ULD7  
 ID Q9ULD7 PRELIMINARY; PRT; 1884 AA.  
 AC Q9ULD7;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DE 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 GN Hypothetical protein KIAA1283 (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=20039619; PubMed=10574462;  
 RA Nagae T., Ishikawa K., Kikuno R., Hiroawa M., Nomura N., Ohara O.,  
 RT "Prediction of the coding sequences of unidentified human genes. XV.  
 RT The complete sequences of 100 new cDNA clones from brain which code  
 RT for large proteins in vitro.";  
 RL DNA Res. 6:337-345 (1999).  
 DR EMBL; AB033109; BAA86597.1; -.  
 DR HSSP; P01024; IC3D.  
 DR InterPro; IPR002890; A2M\_N.  
 DR InterPro; IPR002350; kazal.  
 DR InterPro; IPR001599; Macrogloblina2.  
 DR Pfam; PF00207; A2M; 1.  
 DR Pfam; PF01835; A2M; 1.  
 DR Pfam; PF00050; kazal; 1.  
 DR SMART; SM00280; KAZAL; 1.

DR PROSITE: PS00477; ALPHA\_2\_MACROGLOBULIN; 1.  
 KW Hypothetical protein.  
 FT NON TER  
 SQ SEQUENCE 1884 AA; 206524 MW; D3C078F26B951D1A CRC64;  
 Query Match 21.0%; Score 1542.5; DB 4; Length 1884;  
 Best Local Similarity 27.4%; Pred. No. 7.1e-85;  
 Matches 494; Conservative 257; Mismatches 555; Indels 497; Gaps 57;  
 5 PLTAAHLCLVC-----TALAVALAGPRPLVTAGIIRPGGNVTIIGELLEHCQSQTVA 60  
 7 PLPLPLLLLSARDGVRAPAPAG--YLLAPSVFRAGVEEYVITIP-NSREYTVQA 63  
 61 ELKLTASNLTVSLAEAG--VEFKGSFKTLPLPLNSADEIYELRVTGR---TODEILF 116  
 64 QVLAQGE-----PVVQSGAIIIDKGTIK-LKPTGLAGQA-----LKVWGRMGQAEGEPLF 114  
 117 SNSTRLSFETKRISVITQTDKALKYRKQVYKFRIVTLFSDFKPKYKTSNLILKDKPSNLI 176  
 115 HNOTSVTVDGRGASVFIQTDKPYRRQPHVLLSIFTVSPMLRPNVEKLEAYIILDPGSHM 174  
 177 QQWLS--QGSDLGISTFQLSHLLGDMSIQVAVNDQTYQSFQVSEYVLPRFEYTLQ 235  
 175 IERHLKPPCCGTTNNSPLSDQPLGEWTFEVMQGHAYNNSFEVQKYLPRFELLIDP 234  
 236 PLVCSMNSKHLNLTAKTYGKRVKQDYV--LTFLPLSFGWK--KNITKTFKINGSAN 291  
 235 PRIOGLDACEITVARATFGKRVAGALTINMTVAGVGVSHVEGRPVLRATKIKIGSD 294  
 292 FSPNDEEMKNVWDSNGLEBYLDLSP---GPEILLTVTESVTGISRNSTVNEFKOH 347  
 295 FDIQVADM-----IPADVBEHGRGVSIWAMVT--SVDSQO----- 329  
 348 DYIIIEFDYTV-----LKPSLNFATKVTIRADNOCLTLEBRNNVYI 391  
 330 ---VAADDSTPVQRQVDIRYSKOTRKQKPPGLAYGVKELSLPDSPP-----AEGTVV 380  
 392 TVTORVYTEWSSNGNOQMEAVQKIN--YT---VPQSGTFKLEFPLEDSSE---LOLK 444  
 381 QI-----KAEILTPDNITYTSEVVSQRLGVGEIYSITPSAOHWLETK 423  
 445 AVFLGSK-----SSMAVHSLFKSPSKTYIQLKTRDENIKVGSFELVVSQNKRLK-E 495  
 424 VMALNCKPVGAQYLPYSLISGWY--SPSQCYLQLOPSPHPLQGEAEVYSVSTCQNT 482  
 496 LSTMVSRGQVAVGQO-----N 513  
 483 LYEYVARGNIVLSGOOPAHHTTOQRSKRAPALEKPIRLTHLSETPPPAPAEAVDVCT 542  
 514 STWFSLTPENSWTPKACVIVYIEDDEGIISDYLKI PVOLVFNKIKLWMSKVKAERSBK 573  
 543 SLHLAVTP--SWPFLRLFLVYRENGEYVADSLQRAVEFFENQSVIYSANETOPGEV 600  
 574 VSLRISVTPDSIVGIVAVDKSVNLNANSDITMEVNVHLELYNTG----- 620  
 601 VDLRIIPAAR-GSCVCAAVDKSVYLLRSGRPLTPAVQFQLEBDYVDSDFGVSREDPFW 659  
 621 -----YIIGMFNNSPAVQECGLAVLTID-ANITKQYITGVYDINAETAEAR 663  
 660 MAGLTAQRRRRSVFPWPGITKDSGPAFLETGLVMTDRVSLNHRQDGLY--TDEAVP 717  
 664 FMEENGHIVDIDFSLGSSPH-----VRKHPEFTWIMLDNMGYRIYQ--EFEYTVPD 715  
 718 ARQPHIGSLV-----AVAPSRHPPTREKAKRTTFPEFTWIMHCLNIDPSGEGSLVAVPD 772  
 716 SITSWATGVISEDGLGTLTPVELQAFQPPFIPLNPSYVIRGEERFALETITPNYK 775  
 773 SITSWGEAVVALSTQGLGI-AEPSLILKTFKPEFVDFMLPALIRGEQKIPLSVYVNWG 831  
 776 DATEV--KYIIESSDKFDILMTISSEINATGH-----QOTLAVPSDGAIVLFPPIPTH 827  
 832 TCAEVYMKLSVPGKIQF-----VGHPEKRVHTYKMKCVAPGEAEPIWVLSFSDL 880

QY 828 GEIPITVALS-----PTASDAVYQOMILVKAEGIEKYSQ 862  
 DB 881 GLNNTITAKALAYGDTNCCROGRSSKHPEENHADRVPVIGVDHVRSSVWVEAGVPRAVY 940  
 QY 863 SIL-----LDLT-----DNRL----- 873  
 DB 941 SAFFCPSRVRHISTPENKYEFOYVQRPLRLTFEDVAVRANDARVALSSGPODTAGMIEIV 1000  
 QY 874 -----QSTLKLTSF----- 882  
 DB 1001 LGHONTSMWISTSKMGEPVASANTAKILSWDEFFRTWISWRGGLIQVGHGPEPNSVYI 1060  
 QY 883 -----SFPNTVYGSERVOIT 898  
 DB 1061 VAWTLPRPEVQFIFSTNGWGMGEFRIMRKGEVDESSEAFTLGVPHGALPGSRATAS 1120  
 QY 889 AIGDVLGSLINGLASLIMPPYGCGEQNNINFAPIYIIDIYLTKKKQLTQNTKEKALSFR 958  
 DB 1121 IIGDWMGPTLNLNLRLPGCCGEQNNIHAPVNVFLKYQIQQLSPEVERETDYLV 1180  
 QY 959 OGQRELLYOREDSFSAFGNYDPBSGTWLSAFVLRGFLPADPYIDQNVLIHRTYTLK 1018  
 DB 1181 QGYQRQLTYKQODGSYSAFGERDASGSMWLTAFVLSFAQARSFLVDPRELAANKSWIT 1240  
 QY 1019 GHQKNGEFMDPGRVYHSELQGNKSPVTLTAIYVTSLL--GYRKYPNIDVQESIHFL 1076  
 DB 1241 QQQQADGSFLAVGRVLANQDQGGIHGIVPLTRAYVVALLETGTASEERGSTDAKHLE 1300  
 QY 1077 SEFSRGISDNTYALITYALSSVGSFPAKEALNMLTWBADEGGMQVWSSSKLSDSQ 1136  
 DB 1301 SAAPLAM-DPYSCALTYALTILRSPPAPEALRLKLSLAIMRDVTHW-----SLSNSMD 1354  
 QY 1137 PR-----SLDIEVAVALLSHPLQFQFSEGPIMRWLSRQNSIGFAS 1180  
 DB 1355 VDKGTFLSFSDPVSQSVASAEVEMTAYALITYLLGDVAALPVVKWLSQQRNALGGS 1414  
 QY 1181 TQDTTVALKALSEFAL-----NMTERTNIQVTVG--PSSPS 1216  
 DB 1415 TQDTTVALQALAEVALISYAGINILTVSLASTNDQYTFELHRTNQVLOTAAIPLSLPT 1474  
 QY 1217 PLAVVQMAVNIISANGFPAICQNLVYVYVK---ASGSSRRRSRSTIONG----- 1262  
 DB 1475 GLFV-----SAKGDCCLMQIDVTYVNDPVAKPAFQLVLSIQEBEAGRPPEMDAS 1526  
 QY 1263 -----AFDLDVAVKENDDLNHYDLNCTCSFSGPGRSGMALMEVNLISGFMPWS 1311  
 DB 1527 AAEGRGDMPPADDDPPADQHOEYK--VMLEVCTRMHLAASSNMAVLEVLISGFRADI 1585  
 QY 1312 BAISLSEI-----VKVEYDHGKLANLYLDSV--NETQFCVNIIPAVRFFKVSNTQDASVS 1363  
 DB 1586 E--SLBQLLDKMKMGKRYEVAGRVLFYFPEIERSCLTCVRFRLRCVCYGRISALPVS 1643  
 QY 1364 IVDYVPPRQAVRSYNSSEV-----KLSGCDLQSDVQ-----GCRPECDGASGSHHS 1410  
 DB 1644 VDYDYEPAFEAATRFYVNSTHSPALRELCAGPCACNEVERAPARGGWMFGESGPAVAPBEG 1703  
 QY 1411 SVT 1413  
 DB 1704 AAI 1706  
 RESULT 14  
 ID 0812J3 PRELIMINARY; PRT; 1885 AA.  
 AC 0812J3;  
 DT 01-MAR-2003 (Tremblrel. 23, Created)  
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)  
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
 DE Alpha-2 macroglobulin family protein VIP.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 NCBI\_Taxid=9606;

[1]  
 RN SEQUENCE FROM N.A.  
 RP Li Z., Wu X., Engvall E.;  
 RA "Cloning and expression of VIP, a novel alpha 2 macroglobulin family  
 RT member";  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY101765; AA050084.1; -;  
 SQ SEQUENCE 1885 AA; 206655 MW; 3FB48551FBF326D CRC64;

Query Match 21.0%; Score 1542.5; DB 4; Length 1885;  
 Best Local Similarity 27.4%; Pred. No. 7.1e-85;  
 Matches 494; Conservative 257; Mismatches 555; Indels 497; Gaps 57;

5 PLTAAHLICVC---TAALAVAGPRELVTAPGIIIRPGANVTIGVELLEHPSQVTVKA 60  
 8 PLPLLLLLLSARDGVRAQAPAG--YLIAAPSVFRAQGEVSVITF-NSPREVTVQA 64  
 61 ELKTAASNLTVSLVLEAGS-VPEKGSFKTLPLPSLPLNSADEIYELATYGR--TODEILF 116  
 65 QLVAGGR---PVVQSGAILDKGTIK-LKVPITGLRGA---LLKVGRCMQAEGPLF 115  
 117 SNSTRLPSETRKISVFIOITDKALKYKPKQEVKFRIVTLESDFKPKYKSLNLIKDKPSNLI 176  
 116 HNOISVTVDGASVFIOITDKPVTRPQRRLISIFTVSPNLKPVNEKLEAYILDPKGRM 175  
 177 QQMLS-QQSDLGIVSKTFOLSHPLGDMISIOVQVNDQTYXSFQVSEVYLPKFEVLTQT 235  
 176 IEMHLKPFCCGINTMSFPLSDQVLEGEMFIFVEMOGHAYKSFVQKVLPKFEILLDP 235  
 236 PLYSKMSKHLNGITTAQYVYGVKPKVDVT--LFLPLSPFGKK--KNITKTFKINGSAN 291  
 236 PRYIODDACEGTGTRARATFGKRPVAGALTNMTVNGVGVSHVGPVLATTKTILISQD 295  
 292 FSPFDEEMKVMDSNGISEYLDLSP---GPRLEILTVESTVGTISRNVSTVFPKQH 347  
 296 FDCIVRDM-----IPADVPEHFRGVSIAMAVT-SVDGQQ----- 330  
 348 DYIIEFPDYTV-----LKPSLFTATVKTTRADGNQLTLEERRNNVVI 391  
 331 ---VAFDSTFVQQLVDIRSKDTRKQFKGLAVGVKELSYDGSB---AEGTV 381  
 392 TTTORNTYEWYSGNSGNQKMEAVOKIN-YT--VPSGTFKIEPILLESSE---LQLK 444  
 382 QI-----KAELTPKDNITYSEVSGGLVGFEPISPTSAQVWLETK 424  
 445 AYPFGSK-----SSMAVHSLFKSPKTYIOKTDENIKYKSPRELVYSGNKRK-E 495  
 425 VMAINGKPVGAQVLPSTYLSIGSWY-SPSQCYLQLOPSPHLOVEBAVFSYKSTPCNFT 483  
 496 LSYVWVSRGQLVAVGQ-----N 513  
 484 LYEVAARAGNIVLSGQCPRAHTTQORSKRAALKEKPIRLTHLSTEPPEAEAVDYCVT 543  
 514 STMSLTPENSWTRKACVIVYIIBDGEIISDVLIKIPQVLKFNKIKLYMSKYAEPSEK 573  
 544 SLHLATVP--SMVPLGRLLVFYRENGEGVADSLQFAVETFEFQVSVTASANETQGEV 601  
 574 VSLFISTVQPSIYIVAVDKSVNLMAASNDIMENYVHELELYNTG----- 620  
 602 VDLRIAR-SSCVAVADVSVYLLRSGRFLTPAQVFOELEDDVDVSDSFGVSRDEDPFW 660  
 621 -----YILGMFNNSFAVFOECGLMWLTD--ANTLKDYIDGYVDAVYAE 663  
 661 MAGLTAQRRRRSSVFPWPMGITKDSGFAFTETGLVMTDRVSLNHRDDGLY--TDAV 718  
 719 AFQHTGSLV-----AVAPSRHPRTKRTKRTPEPTMIMCLNISPSGSGTSLVAVPD 773  
 716 SITSWATGFIYSEDGLGTLTTVELOAFOPFIFLNLPSVSVRGEPALETITNYLK 775  
 774 SITSMVBAVALSTSGIGI-AEPSLTKTFKPFVDFMLPALIRGEQVKPLISVYVYMG 832

QY 776 DATEV--KVIIKSKDFDILMTSSEINATGH-----QOTLVPSDEGATVLPFIRPTHL 827  
 DB 833 TCAEVVMKLSVPRKGIQF-----VGHPEKRHYTKMKCVAPGEAEPIMVWLSFSDL 881  
 QY 828 GEIPIYTVALS-----PTASDATQWILYKAEIGKYSQ 862  
 DB 882 GLNNITAKALAYADTNCRRDGRSSKHPENHADRAVPIGVDRRSVWEAEGVPRAYTY 941  
 QY 863 SIL-----LDLT-----DNL----- 873  
 DB 942 SAFFCSERVHISTPKNEYQYQRLRFLRPVAVRAHNDARVALSSGQDTAGMIETV 1001  
 QY 874 -----OSTLKLTSF----- 882  
 1002 LGGHQWTRSWISTSKMGEPVASAHAKILSMDFRTFMISWRGLIOWGHGPEPSNESVI 1061  
 DB 883 -----SPPNPTVTSERVQIT 898  
 1062 VANTLPPPEVQIFGISTGMSKGEFRIWKKEVDESSEAFLLGVPHGAI PGESEATAS 1121  
 QY 899 AIGDVLGPSINGLASHLRMPYGCQEQNMINFANITYILDYTKKKOLDNLKELKALSFMR 958  
 DB 1122 IIGDWGPTLNLNLNLRLPFGCGEQNMIFANVAVLKLTQKQSLSPVEVERETDYLV 1181  
 QY 959 QGYORELLYOREDGSFAFGNNDPSSGTWLSAFVLRCPLEADPYIDIDONVLRHTYTWLK 1018  
 DB 1182 OGVOQLTVYKQGSYSAFGERDASGSMWLTAVLKSFAQARAFIIVDPRELAASKWIT 1241  
 QY 1019 GHQKSGEFMDPRVHVHSELQGNKSPVLTATVYTSLL--GYRKQPNIDVQESIHFLF 1076  
 DB 1242 QQQQADGSGFLAVGVNKKIDOGIHHIIVPLTAAVVVALLETGTASEERSTDKAHFLE 1301  
 QY 1077 SEFSRGISDNYLTALITLYALSVSPKAXELNMLTMRABEQSGMOPWVSSEKLSDSMQ 1136  
 DB 1302 SAAPLPM-DPYSCALTYTTLIRSPAEPALAKLSLAIMRGVTHW-----SLSNSMD 1355  
 QY 1137 PR-----SIDIEVAAVALLSHFLOFQTSBGLPIMKLSRQNSLGGFAS 1180  
 DB 1356 VDKGTFLSFSDRVQSQSVSAEVEMTAVALLTYLLGDVAALAPVVKWLSQQRNALGFS 1415  
 QY 1181 TQDTYALKLSLSEFAL-----MTRRTNIOQTVTG--PSSPS 1216  
 DB 1416 TQDTYALQALAYALLSYAGINTVSLASTNLDYQETFEELHRTQKVLQTAIPLSLPT 1475  
 QY 1217 PLAVVPMVANNISANFGFALCOLNVVYVYK--ASGSSRRRSRSIONOE----- 1262  
 DB 1476 GLFV-----SAKEDCCIMQIDVTYVNPVPAVAPQVLVSLQEPPEQGRPPMPAS 1527  
 QY 1263 -----AFDLVAVKENDDLNHYDLNVCTSPSGRSGMALMEVNLISGFMPVS 1311  
 DB 1528 AAGSGRDMPPADDDPADQHHQYK-VMLBVCTRMLHAGSSNMAVLEVPILSGFRADI 1586  
 QY 1312 EALSISLET-----YKQVEYDHGKANTLYDSV-NETQPCVNI PAVRNFRVSTQDASVS 1363  
 DB 1587 E--SLQQLLDKRMKMGKRYEAVAGRVLFFYFDELPSRCLTCVFRALRECVGRTSALPVS 1644  
 QY 1364 IVDYEPBRAQVAVSYNSEV-----KLSSCDLCSQVQ-----GCRPCEDGASGSHHS 1410  
 DB 1645 VDYIEPAFAATFVNVSTHSPILARELCAGPACNEVERAPARGPMPFSGESGAVAPBEG 1704  
 QY 1411 SVI 1413  
 DB 1705 AAI 1707

RESULT 15  
 ID 060486 PRELIMINARY; PRT; 1476 AA.  
 AC 060486;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, last annotation update)  
 DE Alpha-macroglobulin precursor.

OS Cavia porcellus (Guinea pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Hystriognathi; Caviidae; Cavia.  
OX NCBI\_TaxID=10141;  
RN  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97164019; PubMed=9010766;  
RA Iwasaki H., Suzuki Y., Sinohara H.;  
RT "Cloning and sequencing of cDNAs encoding plasma alpha-macroglobulin  
and murinoglobulin from guinea pig: Implications for molecular  
evolution of alpha-macroglobulin family.";  
RT Biochem. 120:1167-1175(1996).  
RL EMBL; D84336; BAA12316.1; --  
DR HSSP; P01023; 1BV8.  
DR InterPro; IPR002890; A2M\_N.  
DR InterPro; IPR001599; MacrogloblnA2.  
DR InterPro; IPR005531; Tonb\_boxc.  
DR Pfam; PF00207; A2M\_1.  
DR Pfam; PF01835; A2M\_N; 1.  
DR PROSITE; PS00477; ALPHA\_2\_MACROGLOBULIN; 1.  
DR PROSITE; PS00430; Tonb\_DEPENDENT\_REC\_1; 1.  
KM Signal.  
FT CHAIN 1 23 POTENTIAL.  
FT SIGNAL 24 1476 POTENTIAL.  
SQ SEQUENCE 1476 AA; 163614 MW; 0550693A068DC5C5 CRC64;  
  
Query Match 20.3%; Score 1491; DB 11; Length 1476;  
Best local Similarity 28.8%; Pred. No. 6.7e-82;  
Matches 451; Conservative 292; Mismatches 524; Indels 300; Gaps 60;  
  
QY 4 PRLTAHLVCCTAALAVAGPRPLVTAAGIRPGANVTIGELLEHGPSQVTVARL 62  
DB 11 PVL-----LLILAADASISGKPGQVWVIVSLHSGPEKICL-LTQINETVTVASLD 65  
QY 63 -LKTASNLVSVLEAGVPEKSGFKTLTLPSPLENSADELYELRVGTODELPENSTR 121  
DB 66 TIRENSLPMNVAEKDLPCVAF---TVQSPYPAVWELTVKVEGPTHG--FRSRKT 119  
QY 122 LSPETRISVFIQTDKALYKPKQEVKPRVTLFSDPKPYKTSLN-----ILIKDPKSNLI 176  
DB 120 VLAKSDSLVFTVDPRYKPGQTVAKRVSLDENRFP-----LNEIFPLIFIDPRGANV 175  
QY 177 QQWLQQOQSDLGIVISKTFQSLSHPLDWSIQV-QVNDQTYQSGQVSEVYLPRFEVTLQT 235  
DB 176 MQQNLKLERGLQLQSLFSLSEBPLDYSVYVHKEGGRHHSFTVEFVLPREFVQVM 235  
QY 236 PLYCSNMSKHLNCTITAKTYGKPVYGVDTLTFPL-----PLSFWKKKNIYTFE-KING 288  
DB 236 PKKITLIEQEFVTVSGCRYTYGKPVGNITMSICRYNNPNSACLSEB--SRAFCCKYNQ 292  
QY 289 SAN-----FSFNDEBK-----NVMDSNGLSEYLDLSSRGPVEILLTYT 328  
DB 293 QANSQCCFIQOVKTNFQLRKEKYEMLRVEAKIREEGTG---VOLGTGSEITATYT 348  
QY 329 ESVTGISRVSTNVFQKHDIIEFPDYTVLKRPSLNFATVAVKVTADGNQLLEERRNN 388  
DB 349 K-----LSFVKVDSYVRGVPFGQVRL--VDKKNVPMHKM-- 383  
QY 389 VATTQRTYEWSSGNSGNQMEAVQ-KINTYVQSGTFKIEPILDESS---ELQL 443  
DB 384 ITITAEAVY---HSNATTDENGLVQFSINTNMIQTSINLQVHGKSTNCYDQWML 438  
QY 444 KAYFLGSKSMVAHSLFKPSKTYIOLKT-----RDENIKVSGPFELVYSGNRLKE 495  
DB 439 EANEAGSHTAANAFTSL-----SRSFVLEPQLGKLPCHQOTQFAA---HYILKG-QELKE 489  
QY 496 L--SYNVSQGLVAVG-----KONST-----MFSLTPEMSWTPKACVIVYIEDDEGIT 543  
DB 490 LVFVYVYIMAKGGLVQSGTVYLVSEQNTKCHPSVPSVLESDLARVAVLIYALIPSGEIT 549  
QY 544 SDVLKIPVQLVFNKNIKLYMSKVABPSEKVSRLISVTPDSTIVGIVAVDKSVNL----- 598  
DB 550 AUSAQYNVENCIDNKNVNLSPSEGSILPASKTHLRVTAS-POSTICALRAVDQSVLAKRPA 608

QY 599 -MNAAN-----DIT-----MENVHELELYNTGYLGMF-----MNSFA 631  
DB 609 VLSASSVYALLPVKDLTPGRLGLGOEENDECEVSLYNT-YIDGLYPERPINKEKMYG 667  
QY 632 VFQECGLVWLDANLTKOYIDGVNDAEYAEERFEMEENEGHIVDHFSLGSSPH----- 685  
DB 668 FLKDWGLKVFNTKLIKPOLCAHVQKFEVPTMAVSYSESS-----SFRSGPRVPVAVG 720  
QY 721 IATYSEPPKETRYSPETIWDKVIDSSG---VAEVEVTPDTITTEWAKAGACLSND 777  
DB 721 IGLGLTTTPVELOAFQPFPIFLNLPYSVIRGSEFALETIENYLDKATEVKYIIEKSPK 790  
QY 778 TGLGLSPF-ASIRAQPFPELTMYSVIRGEAFILKATVANYLPDCLRISHLEASRKF 836  
DB 686 -----VRGHPEPTW---LDTNMGRYQREFVTPPSISWATGVISD 730  
QY 791 DILMTSSEIMATGHQDTLLVPSBDGATVLPFIRPTHLGEIPTVTA----- 836  
DB 837 LAEPRAKE-----QESYCVCGNERQTVSWVTPKSLGNVNFVSAEALSESSELCNKEK 890  
QY 837 LSPT--ASDAVYQMLLVKAEIGEKYSOSILDLTDNLOSTLKTLSFPENVTGSEK 894  
DB 891 VPTYGKQDTIILKPLVEPEGIEKETWTSIRVSDTVS---EKHLDELPSNVYIDSBAR 947  
QY 895 VQITAGVLPSPINGLASLIRMPYCGEGRNMINPAPNITLYDYTKKKOJLTDNLKEXAL 954  
DB 948 ATVSLIGLILGSAMONIQNLQMPYCGEONWLPAPNITLYDYTNELQOLTPDKSKAT 1007  
QY 955 SFMRQYORELLYOREDSFSAFG-NY-DPSGSTWLSAFVLRCFLEADPYDIDQNVLAR 1012  
DB 1008 SYLSGYQRLNKKRQDSYSTFGENYRGQGNWLTAFVLKTPSQARKYIFIDAHITQ 1067  
QY 1013 TTYTLKQHOKSNGEWDGRIYHSELQCKNSPVTLTIVYTSLSIGYRKQPNIDVQSI 1072  
DB 1068 ALSWLSOKQKONGCPSWSSGSLNNNAIKQGVDEISLSAYITIALLEMSLPDPHPVVRNAL 1127  
QY 1073 HPLES-----EFSRGISDNYTIALITYALSVSQ-PAKKEALNMLTRABEGGMQPMV 1125  
DB 1128 FCLBSAMKSAKEGTG-SHYITKALAVAFALAGQERKKEILKSLDEBGVEDNSLHMA 1186  
QY 1126 SSES-KLSDSW-----QPSRLDIEVAAYALLSHILOFQTSSEGP-----IMRWLSR 1170  
DB 1187 RQKRPVBSGLFKSQAPSABEVMTSYVLAVL-----TARAPTEPDLTSATDYNWTK 1242  
QY 1171 QNSLIGFASLQDTTVALKALSEFAA--LMTERTNIOVTVGSSSP-----LAV 1220  
DB 1243 QONSHGYSSTODTVVALHALSKYAAATFRTTEKA-AQVITKSSGTFSTNFEVNHNRL 1301  
QY 1221 VQPMV-----NISANGPGALCOLVVVNV-----KASSSSRRR 1256  
DB 1302 LQOVSLPVTSDSYITTVTGBGVYIQTSLKTNVSEKCTPPPALAEFTVPOACDPKANT 1361  
QY 1257 SIONQEAFLDVAVKENKDINHVDLNVCTSPSGRSGMALMEVNLISGF--MVPs-EA 1313  
DB 1362 SFO-----ISLNVSYGSR-VSNMAIVDKMWSGFIPLKPIYKN 1400  
QY 1314 ISLSTVKKVEYDHGKLLYIDSVNETQFCNIPAVRNFKVSNODASVIVDYEPERQ 1373  
DB 1401 LEKSEHISRTESNNHVLIDYDKVSNQTLSSFFVQDIEVRDLKPAIKKYDYETNEF 1460  
QY 1374 AVRSYNS 1380  
DB 1461 AAEYHA 1467

Search completed: January 15, 2004, 18:14:03  
Job time : 64 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: January 17, 2004, 00:12:55 ; Search time 5629 Seconds  
(without alignment)  
6165,712 Million cell updates/sec

Title: US-10-020-095-4  
Perfect score: 7348  
Sequence: 1 MGGPPLTAHLICVCTAL.....HSSVIFCFKLYFMELML 1428

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlp  
-O=/cgn2.1/USPTO.spool.P/US10020095/runat.15012004.163822.18713/app\_query.fasta.1.1607  
-DB=BST -QFMT=fastap -SUFFIX=rcs -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR SCORE=DCT -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pcio -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10020095 @CGN 1.1 4337 @runat.15012004.163822.18713 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEBUFFER -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGJOS  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hlc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hlc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrt1:\*  
28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5343	72.7	4485	11 AK029247	AK029247 Mus muscu
2	5304	72.2	4725	11 AK036799	AK036799 Mus muscu
3	1480	20.1	4641	11 BC044072	BC044072 Xenopus 1
4	1461	19.9	941	13 BX391819	BX391819 BX391819
5	1406	19.1	1201	9 AL554092	AL554092 AL554092
6	1405.5	19.1	1201	9 AL554092	AL554092 AL554092
7	1280	17.4	879	13 BU517328	BU517328 Mus muscu
8	1117	15.2	667	12 BQ028040	BQ028040 UI-H-CO0
9	1022	13.9	1201	9 AL576908	AL576908 AL576908
10	1007	13.7	1495	11 AK042169	AK042169 Mus muscu
11	1006	13.7	978	13 BX386457	BX386457 BX386457
12	979.5	13.3	836	10 EG173799	EG173799 EG173799
13	972.5	13.2	638	13 BU624356	BU624356 UI-H-RG1
14	941	12.8	648	10 BB613975	BB613975 BB613975
15	912	12.4	634	9 AA418644	AA418644 zvc3f07.r
16	902.5	12.3	613	14 CB158337	CB158337 K-BST0217
17	895	12.2	919	13 BX370169	BX370169 BX370169
18	870	11.8	988	13 BX386456	BX386456 BX386456
19	849.5	11.6	586	9 AA284110	AA284110 zc36a07.T
20	833	11.3	473	13 BX279537	BX279537 BX279537
21	831	11.3	537	14 CA417651	CA417651 UI-H-FE0
22	830	11.3	515	13 BX472602	BX472602 DKFZP686C
23	829	11.3	544	10 BE143588	BE143588 MR0-FR016
24	759	10.3	853	10 BE573834	BE573834 601331813
25	750.5	10.2	835	14 CD242054	CD242054 AGENCOURT
26	746	10.2	490	12 B1975931	B1975931 484508 MA
27	738	10.0	437	9 AA132252	AA132252 z006h07.r
28	732	10.0	456	9 A1278292	A1278292 q178d07.x
29	731	9.9	480	12 BM537249	BM537249 ha80a02.g
30	692	9.4	684	9 AV245523	AV245523 AV245523
31	687	9.3	806	12 BG919720	BG919720 602822407
32	671	9.1	722	14 CA420610	CA420610 UI-H-FG0
33	661	9.0	511	14 CA893496	CA893496 B0179D10-
34	660	9.0	620	13 BQ265550	BQ265550 NTSC f13
35	649	8.8	512	14 CA566520	CA566520 K0402C12
36	645	8.8	395	12 BG983552	BG983552 UI-H-CM006
37	636	8.7	717	13 BQ775821	BQ775821 UI-H-FH0-
38	635	8.6	635	14 CA429478	CA429478 UI-H-FH1-
39	633.5	8.6	673	10 BB639005	BB639005 BB639005
40	626	8.5	476	13 BQ551040	BQ551040 H4067H10-
41	614	8.4	687	13 BQ773484	BQ773484 UI-H-FE0-
42	612.5	8.3	419	14 CB158287	CB158287 K-EST0217
43	610.5	8.3	432	13 BU434307	BU434307 603836006
44	609	8.3	431	10 BE006130	BE006130 RC4-BN012
45	597	8.1	652	12 BQ020033	BQ020033 UI-H-ED0-

#### ALIGNMENTS

RESULT 1  
LOCUS AK029247 4485 bp mRNA linear HTC 05-DEC-2002  
DEFINITION Mus musculus 0 day neonate head cDNA, RIKEN full-length enriched  
library, clone:4831440K17 product:weakly similar to  
alpha-2-macroglobulin (fragments) [limulus polyphemus], full insert  
sequence.  
ACCESSION AK029247  
VERSION AK029247.1 GI:26081244  
KEYWORDS HTC; CAP trapper.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCES

1  
Mammalia; Euteria; Rodentia; Sciurognathi; Muridae; Murinae; Mus  
Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)

2  
10349666  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)

3  
20499374  
11042159  
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,  
Kono, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,  
Sun, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,  
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kaishogi, K.,  
Yuiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, K.,  
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J.,  
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multipicillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)

4  
20530913  
11076861  
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,  
Aizawa, T., Hara, A., Fukunishi, Y., Kono, H., Adachi, J., Fukuda, S.,  
Arikawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamada, I.,  
Saito, T., Okazaki, Y., Gojohori, T., Bono, H., Kaikawa, T., Saito, R.,  
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,  
Fleischmann, W., Gaasterland, T., Gissi, C., Kling, B., Kochwa, H.,  
Kuehl, P., Lewis, S., Matsumoto, Y., Nikaido, I., Pesole, G.,  
Quackenbush, J., Schriml, L.M., Stabili, F., Suzuki, R., Tomita, M.,  
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,  
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,  
Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C.,  
Fletcher, C., Fujita, M., Gariboldi, M., Gusttinch, S., Hill, D.,  
Hochmann, C., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,  
Marchionni, L., Mashima, J., Mazzarelli, J., Momberts, P., Nordone, P.,  
Ring, B., Rindwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,  
Sato, K., Schombach, C., Seyer, T., Shibata, Y., Storch, K.F., Suzuki, H.,  
Toyo-Oka, K., Wang, K.H., Weitz, C., Whitaker, C., Wilming, Y.,  
Wynshaw-Borle, A., Yoshida, K., Hasegawa, Y., Kawai, J., Kontsuki, S.  
and Hayashizaki, Y.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409 (6821), 685-690 (2001)

5  
21085660  
11217851  
The FANTOM Consortium and the RIKEN Genome Exploration Research  
Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation  
of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

JOURNAL  
MEDLINE  
PUBMED  
REFERENCES  
AUTHORS  
TITLE

6 (bases 1 to 4485)  
Adachi, J., Aizawa, K., Akimura, T., Arikawa, T., Bono, H., Carninci, P.,  
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,  
Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T.,  
Hori, F., Imocani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,  
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M.,  
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,  
Nakamura, M., Nishi, K., Nomura, K., Nunazaki, R., Ohno, M., Ohsato, N.,  
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,  
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,  
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,  
Takeda, Y., Tanaka, T., Tomaru, A., Toyo, T., Yasunishi, A.,  
Muramatsu, M. and Hayashizaki, Y.  
Direct Submission  
Submitted (16-JUL-2001) Yoshinide Hayashizaki, The Institute of

COMMENT

Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama Kanagawa 230-0045, Japan (E-mail: genome-res@gs.c.riken.go.jp, URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9216, Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: <http://genome.gsc.riken.go.jp/>  
 URL: <http://fantom.gsc.riken.go.jp/>  
 Location/Qualifiers

FEATURES

source

1. .4485

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C57BL/6J"

/db\_xref="PANTOM,DB:4831440K17"

/db\_xref="taxon:10090"

/clone="4831440K17"

/tissue\_type="head"

/clone\_lib="RIKEN full-length enriched mouse cDNA library"

/dev\_stage="0 day neonate"

misc\_feature

112..4439

/note="putative weakly similar to alpha-2-macroglobulin (fragments) [Imlulius polyphemus] (PIR/A36260, evidence: FASTY, 51.8%id, 76.7%length, match=166)"

BASE COUNT 1277 a 970 c 1010 g 1228 t

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 4485

Score: 5343.00 Matches: 1036

Percent Similarity: 83.70% Conservative: 176

Best Local Similarity: 71.55% Mismatches: 206

Query Match: 72.71% Indels: 31

Gaps: 7

US-10-020-095-4 (1-1428) x AK029247 (1-4485)

OY 1 MecGInGlyPProPLeuLeuThraAlaAhiSLeuLeuCyValCyThraAlaLeu 20

Db 112 ATGGCGAGTCGAGAGCTCTCTGAGCGCGCCACCTTCTGCTTGCGCGGTCCGCG 171

OY 21 AlaValAlaPProGlyPProArgPheLeuValThraAlaPProGlyIleIleArgPProGly 40

Db 172 GCC--GCCCGAGGGTCCCGCTTCTGGTGAAGCCCGCAGGAGATATCAAGACCCGAGCA 228

OY 41 AsnValThriIleGlyValGluLeuLeuGluHisCySProSerGlnValThraValLySAla 60

Db 229 AATGGAATATTGGGGTGGATCTCTCTGAAAAATAGCCCCCAGACAGCTCTGTAAGCT 288

OY 61 GluLeuLeuLySThraAlaSerAsnLeuThraValSerValLeuGluAlaGluGlyValPhe 80

Db 289 CAGGTGTTCAAGATAGCTTCCAAACAAATCGAGATCTCATCTCGGAAGCAAGAGACTTT 348

OY 81 GluLySgLySerPheLySThreLeuThreLeuProSerLeuPProLeuAsnSerAlaAspGlu 100

Db 349 CACAGAGGCCATTTCMAAGACTCTCGTTTCCGGCATCACTCTGAGCAATGCAAGATAG 408

OY 101 IleTyrgLueAurGValThrgIYAgrthGlnAspGluIleLeuPheSerAsnSerThr 120

Db 409 ATTATTAGCTACACATTAAGGACAAATCAAGAAATGAGATGATCTATTCTCCCAAGAGACA 468

OY 121 ArgLeuSerPheGluThrLySArgIleSerValPheIleGlnThraSplyValLeuTy 140

Db 469 CGTTTAACATTGAGAGCAAGATATATCTCTCCATTCAGACAGACAAGGCAATTCAC 528

OY 141 LySProLySgInGluValLyPheArgIleValThriLeuPheSerAspPheLySProTy 160

Db 529 AAGCCGAAGCAGAAAGTAACTTTGGGCTGCACACTCTGCTCAAGTCTCAAGCCTTAT 588  
 Qy 161 LysThrSerLeuAsn11Leu11LeuYsaProLysSerAsnLeu11eGlnGlnTrpLeu 180  
 Db 589 AGGACATCGTGGAGCATTTTCACTTAAGGATCCGAAGTCTTAAGTATCAACAGAGGTTT 648  
 Qy 181 SerGlnGlnSerAspLeuGlyVal11SerLysThrPheGlnLeuSerSerHisProIle 200  
 Db 649 TCTCGAAAGGTCAGTCTGGAGTCGTTCCAAACCTTTCAACTATCCCTCAATCCAAATA 708  
 Qy 201 LeuGlyAspTrpSer11eGlnVal11eGlnVal11eAsnAspGlnThrTrpTrpGlnSerPheGln 220  
 Db 709 TTTGGTGAATGGTCACTTCAAGTCAAGTGAATGATCGACATATTTACAGCATTTTACG 768  
 Qy 221 ValSerGlnTrpVal11eLeuProLysPheGlnVal11ThrLeuGlnThrProLeuTrpCysSer 240  
 Db 769 GTCCATGAATATGTTTACCAAAAGTTCCAAAGTCACCGTCGACAGACCCCGCTGATTTGCC 828  
 Qy 241 MetAsnSerLysHisLeuAsnGlyThr11LeuThrAlaLysTrpTrpTrpVal11e 260  
 Db 829 TTGAAGTCTAAACAGTAAATGAGTCCGTCATAGCAAAAGTATACATACGGGAAGCCAGTG 888  
 Qy 261 LysGlyAspVal11ThrLeuThrPheLeuProLeuSerPheTrpGlyLysLysLysAsn11e 280  
 Db 889 AAGGGAAGTCTGTCACTTACATTTTACCTTTGTCTTTTGGGAAAGAAAGAAATATC 948  
 Qy 281 ThrLysThrPheLys11eAsnGlySerAlaAsnPheSerPheAsnAspGlnGlnMetLys 300  
 Db 949 ACAAAAAGTTTGAAGATTAATGATTTGCAAACTTCTCTTTGATTAATTAAGAGTGA 1008  
 Qy 301 AsnValMet-----AspSerSerAsnGlyLeuSerGlnTrpLeuAsp 314  
 Db 1009 AAGGTAATGAATTTGAAACCGCTAACGAGTGTCTCGAAGGAGGATTAAGAAATGTGAC 1068  
 Qy 315 LeuSerSerProGlyProValGln11eLeuThrThrVal11ThrGlnSerVal11ThrGly11e 334  
 Db 1069 CCGTCTTCTTCCGGCGCGCAGAGATTTAGCTACTGTGACGGAATCTCACTCACAGTATC 1128  
 Qy 335 SerArgAsnVal11eSerThrAsnVal11PhePheLysGlnHisAspTrp11eGlnGlnPhe 354  
 Db 1129 TCAGAAATGGCAGACCAATGTGTTTTCAGAGCAACATGATTAACATCTATGAGATTTT 1188  
 Qy 355 AspTrpThrThrVal11eLysPheSerLeuAsnPheThrAlaThrVal11LysVal11ThrArg 374  
 Db 1189 GATTATACAGCTGCTTGAACCATCTCTCAACTTCACAGCCACGTGAAGTACGCCGC 1248  
 Qy 375 AlaAspGlyAsnGlnLeuThrLeuGlnGlnLysArgAsnAsnVal11eThrVal11Thr 394  
 Db 1249 TCTGATGGCAATCAACTAATCTCTGAAGAAATGAATAATGATTTGTCAAGTATGAC 1308  
 Qy 395 GlnArgAsnTrpThrGlnTrpSerGlySerAsnSerGlyAsnGlnLysMetGlnAla 414  
 Db 1309 CAGGGAATAATAC-----ACCTT-GAGAGTCAAGAGATCAGAGAAATGATTAAT 1358  
 Qy 415 ValGlnLys11eAsnTrpThrVal11ProGlnSerGlyThrPheLys11eGlnPheProIle 434  
 Db 1359 ATCCAGACCGTAAATTAATATCTATCCCCCAAAATGAAATCATTAATGAATTTCCCAATC 1418  
 Qy 435 LeuGlnAspSerSerGlnLeuGlnLeuLysAlaTrpPheLeuGlySerLysSerSerMet 454  
 Db 1419 ATGTTCATTTCCGGAGAGCTCAACTGAAGGCTTATTTCTTGAAGTGAACAAGTTCTGTG 1478  
 Qy 455 AlaValHisSerLeuPheLysSerProSerLysThrTrp11eGlnLeuLysThrArgAsp 474  
 Db 1479 ACAGTCCACAGATATTCACCTCTCTAGTAAAGAGTCAATCAACTTAAAAACGAGAAAT 1538  
 Qy 475 GlnAsn11LysVal11GlySerProPheGlnLeuVal11eSerGlyAsnLysArgLeuLys 494  
 Db 1539 GAATATATTAAGAGTGGGTCACCTTTGATTTGATGCTTAAAGTGGCAACAGAAATTCAG 1598  
 Qy 495 GlnLeuSerTrpMetVal11eSerArgGlyGlnLeuVal11eAlaGlyLysGlnAsnSer 514  
 Db 1599 GACTTAAGCTATATGATATATCAAAAGGACAGTGTGCTGACAGCAAAACAAGTTCA 1658

Qy 515 ThrMetPheSerLeuThrProGlnAsnSerTrpThrProLysAlaCysVal11eValTrp 534  
 Db 1659 AGAATCTTCTCTTAAACCAAGAGGCTTCCGTGGGCTCCAAAGGCTGTATATGCAATAT 1718  
 Qy 535 Tyr11eGlnAspArgGlyGln11eLysSerAspVal11eLysVal11eProVal11eGlnLeuVal 554  
 Db 1719 TATATGCAAGAGATGGGGAATTAATATATTTCAAAAATCCCGCTTACAGCTTGT 1778  
 Qy 555 PheLysAsnLys11eLysLeuTrpTrpSerLysVal11eGlnLysProSerGlnLysVal 574  
 Db 1779 TTGAATAATAGTAAAGCTGTTTGGAGTAAACCTTCAAGTCAAGCCATTCGATTAAGTCT 1838  
 Qy 575 SerLeuArg11eSerVal11ThrGlnProAspSer11eVal11eGly11eVal11eVal11eAspLys 594  
 Db 1839 TCTCTCAGAGATCTCTGCGACACAGTCTGATCCCTGCGGGGATTTGATGCTGTGACAAA 1898  
 Qy 595 SerValAsnLeuMetAsnAlaSerAsnAsp11eThrMetGlnAsnVal11eGlnLeu 614  
 Db 1899 AGTGAATCATGATGAAACTCCAAACAGTATTAATGAAGAAACCATGCTCATGAACCTG 1958  
 Qy 615 GlnLeuTrpAsnThrGlyTrpTrpTrpLeuGlyMetPheMetAsnSerPheAlaVal11PheGln 634  
 Db 1959 GAACCTTATTAACAGAAATATTAATTAATGAGATGTTCAATGAACCTCTTCCGCTGATTTCCAG 2018  
 Qy 635 GlnCysGlyLeuTrpVal11eLeuThrAspAlaAsnLeuThrLysAspTrp11eAspGlyVal 654  
 Db 2019 GAATGTGCTCTCGGAGTTCACAGATGCAACCTTTAAGAGATAGCATTTGAAGATC 2078  
 Qy 655 TyrAspAsnAlaGlnTrpAlaGlnArgPheMetGlnGlnGlnGlnGlyHis11eValAsp 674  
 Db 2079 TACGATTAAGAAAGTATTTGAAAGCTTTCAGAGGAAATGAGCAAACTCGGTGAT 2138  
 Qy 675 IleHisAspPheSerLeuGlySerSerProHisValArgLysHisPheProGlnTrp 694  
 Db 2139 TTTGAAGTGTCTTCTTCAATTAAACATGTATCATATGACGAAGAAATTTCCAGAAACATG 2198  
 Qy 695 IleTrpLeuAspTrpAsnMetGlyTrpArgLysLeuTrpGlnGlnPheGlnVal11ThrValPro 714  
 Db 2199 ATTTGGCTAGCGCTCATGAGATCTTAAGATTTTAAGAAATTTGAAGTATCACTGATCCT 2258  
 Qy 715 AspSer11eThrSerTrpVal11eThrGlyPheVal11eSerGlnAspLeuGlyLeuGly 734  
 Db 2259 GACTTCATCATCTTCTGGGTGGCTTGGCTTTTGTCACTCAGAAACCTGGGTTTGGG 2318  
 Qy 735 LeuThrThrTrpProValGlnLeuGlnAlaPheGlnProPhePhe11ePheLeuAsnLeu 754  
 Db 2319 TTAAACAACCTTCCAGCAGACAGACTCAACCTTCCAAACCGTTTCTCTTCTCAATCTT 2378  
 Qy 755 ProTrpSerVal11eArgGlyGlnGlnPheAlaLeuGlnIleThr11ePheAsnTrpLeu 774  
 Db 2379 CCGTACTGTATTAACAGAGGGAAGATTTGCTTTGGAAGTATCACTGCAATTAATTTG 2438  
 Qy 775 LysAspAlaThrGlnVal11eLysSerAspLysPheAsp11eLeuMet 794  
 Db 2439 AAAAGTAAATTAAGTGTGATATCTATTAAGAAAGTGCACACTTTGATTAATTAATG 2498  
 Qy 795 ThrSerSerGln11eAsnAlaThrGlyHisGlnGlnThrLeuLeuVal11eProSerGlnAsp 814  
 Db 2499 ACTTCAATATGACACCAATGAGCACAATATACCGGAAGACTGTTCAAGTTCCACAGGATTAAT 2558  
 Qy 815 GlyAlaThrVal11eLysPheProIleArgProThrHisLeuGlyGlnIleProIleThrVal 834  
 Db 2559 GGGGTCACTTGTGTTTCCATCAAGCCCAAGCAGCTTGGAGAGATTTCCATCACTG 2618  
 Qy 835 ThrAlaLeuSerProThrAlaSerAspAlaVal11ThrGlnMet11eLeuVal11eAlaGln 854  
 Db 2619 ACGGTGCTCAACCACTGCTGTGATGCTGTCAACCAAGCAATTTGAATGAACCTGAA 2678  
 Qy 855 Gly11eGlnLysSerTrpSerGlnSer11eLeuLeuAspLeuThrAspAsnArgLeuGln 874  
 Db 2679 GGGATAGAAAAGTGTATTTGAAAATCTGTATTTGATGATCGACAGATGACAAAGTGA 2738

Oy		875	LeuAlaValSerThrLeuSerPheSerPheProProAsnThrValThylGlySerGluArg	894
Db		2739	AGTAAGCAGCAATCTATACAGATTCTTCCTTCCCTGTACAGCGCATTTGGCAGTGAAAGA	2798
Oy		895	ValGlnIleThrAlaIleGlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeu	914
Db		2799	GTTCCAGATCACAGCAATTGGAGATATCTTGGTGCCTCATCAATAGGCTTATCTTCACTG	2858
Oy		915	IleArgMetProIlyrGlyCyseGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTy	934
Db		2859	ATCCGAAGCCTTATGATGTGGTGMAACAAGAACAGATATATTGGTCCCAAAATAATTAC	2918
Oy		935	IleLeuAspTyrlLeuThylSylValysGlnLeuThrAspAsnLeuLysGluLysValLeu	954
Db		2919	ATTTTGGATTATCTGACTTAACACAGAAACAGCTGACAGTAAATTTAAGAAAAGGCCCTT	2978
Oy		955	SerPheMetArgInglYTyrglnArgGluLeuLeuTyrglnArgGluAspGlySerPhe	974
Db		2979	TCAATATAGAGCAAGGTTATCCAAGAGAGGCTTCTATCAAGAGGAAGATGGCTCCTTC	3038
Oy		975	SerAlaPheGlyAsnTyrlAspProSerGlySerThrTrpleuSerAlaPheValLeuArg	994
Db		3039	AGTGCTTTGGGGACATTCAGACTCTTCTGGAGACCTTGCTGACAGCTTTGTTTTAAGA	3098
Oy		995	CysPheLeuGlnAlaAspProTyrlLeaAspIleAspGlnAsnValLeuHisArgThyr	1014
Db		3099	TGCTTTTCGGAAGCGATTACTATATATATATATGTCAGATGTGTATCACAGAACATAT	3158
Oy		1015	ThrTrpleuLysGlyHISglnLysSerAsnGlyGluPheTrpAspProGlyArgValIle	1034
Db		3159	ACTTGCTTAATGCACATTAAGAAATTCATGTGTAAATTTGGAGCCAGAAAGTAGTATT	3218
Oy		1035	HisSerGluLeuGlnGlyGlyAsnLysSerProValThrLeuThrAlaTyrlIleValThr	1054
Db		3219	CACAGTGAACCTTCAAGGTGGACCACCAAAGCCAGTACACTTACCGCTTAATTGTGACT	3278
Oy		1055	SerLeuLeuGlyTyrlArgLysTyrcIlnProAsnIleAspValGlnGluSerIleHisPhe	1074
Db		3279	TCTGTCTCGGGATACAAAAAGTACGCTTAATATCGATATGTACAAAGTCAATCAAGTTT	3338
Oy		1075	LeuGlnSerGluPheSerArgGlyIleSerAspAsnTyrlThrLeuAlaIleIleThrTy	1094
Db		3339	TTGGAATTGTGAATTCAGCAGAGAGATTTCCGACATTTATACCTTAGCAATTATATCCAT	3398
Oy		1095	AlaLeuSerSerValGlySerProLysAlaLysGlnAlaLeuAsnMetLeuThrTryArg	1114
Db		3399	GCCCTGTCCACAGTGGGAGCCCTTAACAGAGAGAGCGCTTGAATCTGTATGACGGA	3458
Oy		1115	AlaGlnGlnGlnGlyGlyMetGlnPheThrValSerSerGluSerLysLeuSerAspSer	1134
Db		3459	TCAGAAAAGGAAGACACACACAGTTCGTGATCATAGGACCTGCCTCTCGGGTCC	3518
Oy		1135	TyrGlnProArgSerLeuAspIleGlyValAlaAlaIlyrAlaLeuLysSerHisPheLeu	1154
Db		3519	TGGCAGCCGCGCTCCGTGACATCGAAATCGCTGCTTATGCACCTGCTGCGCACCTTG	3578
Oy		1155	GlnPheGlnThrSerGluGlyIleProIleMetArgTrpleuSerArgGlnArgAsnSer	1174
Db		3579	CAC---CATGTGTCTGAGAGGAATCCCGGTATGAGGTGGCTCATCACAAAGAAACAGC	3635
Oy		1175	LeuGlyGlyPheAlaSerThrGlnAspThrThrValAlaLeuLysValLeuSerGluPhe	1194
Db		3636	CTGGAGAGCTTTGTATCCACTCAGGATATCTGTGTGGCCCTGAAAGGCAATTATCTGAATTT	3695
Oy		1195	AlaAlaLeuMetAsnThrGluArgThrAsnIleGlnValThrValThylGlyProSerSer	1214
Db		3696	TCAAGCCTTGTGCACAGAAATATCAATATCAACTGACCGGTACAGGGCCGGCATTC	3755
Oy		1215	Pro-----SerPro	1217
Db		3756	CCCAGATCCATACACTTCCGAATTGACTCCACGAACCTTTTCTCCTTACACAGGAAGAG	3815
Oy		1218	LeuAlaValValGlnProMetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIle	1237

Db	3816	CTTCATGCGCTAGATCCCATCACCGCTTAATGTTTCTCCGATGCGTGGATTTTGCATT	3875		
Qy	1238	CysGlnLeuAenValValTyrAsnValLysAlaSerGlySerSerA-ga-r-ga-r-gSer	1257		
Db	3876	TGCCAGCTTAATGTTGACTTAAGTGAAGAGTTTCAGGTTCTTCTTAATAACGAGATCT	3935		
Qy	1258	ILeGlnGlnGlnValaPheAspLeuAspValaValLysGlnAenLysAspAspLeu	1277		
Db	3936	ACCAAAATACAGAAAGTTTGTGATTAGAGTCATTGTG--AAATAATGAGACGACATTT	3992		
Qy	1278	AsnHisValAspLeuAenValCysThrSerPheSerGlyProGlyAsrSerGlyMetAla	1297		
Db	3993	AGTCACCTGAAATCTGAATGTGTGTGCAAGTCACTGTGGTTCAAGAGACGACGATGTGC	4052		
Qy	1298	LeuMetGlnValaAenLeuLeuSerGlyPheMetValProSerGlnValaIleSerLeuSer	1317		
Db	4053	CTTAATGAGAGTGAACCTTCTCAGTGGCTTTAGTGCATCTTCAGATTCATTTCCCTGAGT	4112		
Qy	1318	GlnThrValLysLysValGlnTyrAspHisGlyLysLeuAenLeuTyrLeuAspSerVal	1337		
Db	4113	GAGACCCCTGAAGAAAGTGAGATAAGTAAGAAAGGAAATCAACTTTATTTAGATTGTA	4172		
Qy	1338	AsnGlnThrGlnPheCysValAsnIleProValaValaArgAsnPheLysValSerAsnThr	1357		
Db	4173	AATGAATCCCAAGTTTGTGTTAATATTCCTCCACTGTGAGACTACAAAGTTTCAATATT	4232		
Qy	1358	GlnAspAlaSerValSerIleValaAspTyrTyrGlnProArgArgGlnAlaValaArgSer	1377		
Db	4233	CGAGATGGTTCAATCTCTGTAATGACTACTATGAAACAAAGAGACAGCGACGTCCGAAC	4292		
Qy	1378	TyrAsnSerGlnValLysLeuSerSerCysAspLeuCysSerAspValGlnGlyCysArg	1397		
Db	4293	TACAACCAACCCAGGGAAGCTGTCTCTCAATGTACTCTCAGTCAGACACC--AACTGACAG	4349		
Qy	1398	ProCysGlnAspArgLysAlaSerGlySerHisHisSerSerValIlePheIlePheCys	1417		
Db	4350	TCACACACGAGCGAGCGACGAGACTCCCTTCGACGTTCTTCGAGCCTTCGTCTTTTGC	4409		
Qy	1418	PheLysLeuLeuTyrPheMetGln	1425		
Db	4410	TCCGTCCTTCTGTACTTTGTGCAA	4433		
RESULT 2	AK036799	4725 bp	mRNA	linear	HTC 05-DEC-2002
LOCUS	AK036799				
DEFINITION		Mus musculus adult testis cDNA, RIKEN full-length enriched library, clone:9930012B15 product:weakly similar to alpha-2-macroglobulin (fragments) [limulus polyphemus], full insert sequence.			
ACCESSION	AK036799				
VERSION	AK036799.1	GI:26085434			
KEYWORDS	HTC, cap trapper.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	1	Carninci, P. and Hayashizaki, Y.			
AUTHORS	1	High-efficiency full-length cDNA cloning			
TITLE	1	High-efficiency full-length cDNA cloning			
JOURNAL	1	Meeth. Enzymol. 303, 19-44 (1999)			
MEDLINE	99279253				
PUBMED	10349636				
REFERENCE	2	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.			
AUTHORS	2	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes			
JOURNAL	2	Genome Res. 10 (10), 1617-1630 (2000)			
MEDLINE	20499374				
PUBMED	11042159				
REFERENCE	3	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,			
AUTHORS	3				

Komno, H., Akiyama, J., Nishi, K., Kitsuina, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, B., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system-384-format  
sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCES  
AUTHORS

20530913  
11076861

4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Komno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batilov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Staudli, F., Suzuki, R., Tomita, M., Quackenbush, J., Schriml, L. M., Staudli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Holtman, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombert, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyo-oka, K., Wang, K. H., Weitz, C., Whitaker, C., Wilming, L., Wyshaw-Borja, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohetsuki, S. and Hayashizaki, Y.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409 (6821), 685-690 (2001)

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCES  
AUTHORS

21085660  
11217851

5 The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCES  
AUTHORS

6 (bases 1 to 4725)  
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hayashizaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirakata, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, Y., Kojima, Y., Kondo, S., Komno, H., Kouda, M., Koyama, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numata, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takada, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.  
Direct Submission

TITLE  
JOURNAL

Submitted (16-JUN-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suenitro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.go.jp/  
URL: http://pantom.gsc.riken.go.jp/  
Location/Qualifiers

FEATURES  
source

1. .4725  
/organism="Mus musculus"

misc\_feature  
/note="putative  
weakly similar to alpha-2-macroglobulin (fragments)  
[Linnaeus polypheumus] (PIR|A36260, evidence: FASTY,  
51.8%ID, 76.7%length, match=168)"  
4707..4712  
/sex="female"  
/clone="9930012E15"  
/tissue\_type="vagina"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="adult"  
97..4429  
/note="putative  
polyA\_signal  
4707..4712  
/note="putative"  
polyA\_site  
4725  
/note="putative"  
BASE COUNT 1334 a 1016 c 1066 g 1309 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 0 Length: 4725  
Score: 5304.00 Matches: 1037  
Percent Similarity: 83.49% Conservative: 177  
Best Local Similarity: 71.32% Mismatches: 202  
Query Match: 72.18% Indels: 38  
DB: 11 Gaps: 8  
US-10-020-095-4 (1-1428) x AK036799 (1-4725)  
QY 1 MetGInGlyProProLeuLeuThraAlaHisIseuLeuCyValCyseThraAlaIleu 20  
Db ATGCCAGTCCGAGGCTCTTAAGCGGCCCACTTCTCTGCTTGGCCGCGCTG 156  
QY 21 AAlaValAlaProGlyProArgPheLeuValThraAlaProGlyIleIleArgProGlyGly 40  
Db GCC--GCCCGAGGCTCCCGCTTCTGCTGAGAGCCAGGAGATCATGACCGGAGCA 213  
QY 41 AsnValThrIleGlyValGluLeuLeuGluHisCyPProSerGlnValThraValysAla 60  
Db AATGTGACTAATGGTGGGTGATCTCTGAAAATAGCCCCCAAGGTCCTGTAAAGCT 273  
QY 61 GluLeuLeuYThrAlaSerAsnLeuThraValSerValLeuGluAlaGluGlyValPhe 80  
Db CAGGTTCAGATRAGCTTCAACAATCGAGATCATCTCGAAGCAGAGAGGTCTT 333  
QY 274 GlnuysGlySerPheLeuThraLeuThraProSerLeuProLeuAsnSerAlaAspGlu 100  
Db CACAGAGGCCATTCAACACTCTCTTCCGACACTACCTCGACAGATGACATGAG 393  
QY 101 IleYrGluLeuArgValThraGlyArgThrGlnAspGluIleLeuPheSerAsnSerThr 120  
Db ATTTATGAGCTACACATTAACGACCAATCAAGATGAGATGATCTCTCAACAGCACA 453  
QY 121 ArgLeuSerPheGlnThraValArgIleSerValPheIleGlnThraAspValAlaLeuYr 140  
Db CGTTTAACATTGAGAGGAGGATGATATCTGCTCTCAATTCAGACAGACAGGATTTAC 513  
QY 141 LysProLysGlnGluValIleValPheArgIleValThraLeuPheSerAspPheLeuProYr 160  
Db AAGCGGAGGAGGAGGATGATTTCCGAGTCTCACTCTGCTGATCTGATCTGATCTAT 573  
QY 161 LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGlnThraLeu 180  
Db AGGACATCCGTCGATTTTCAATTAAGATCCGAAGTATGATGATCAACAGGAGTTT 633  
QY 181 SerGlnGlnSerAspLeuGlyValIleSerYrThrPheGlnLeuSerSerHisProIle 200  
Db TCTCAGAAAGGTGATCTGAGTCTTTCACAACTTTCACATCTCTTAATCAATCA 693  
QY 201 LeuGlyAspTrpSerIleGlnValGlnValAsnAspGlnThrYrYrGlnSerPheGln 220

Db 694 TTGGTGAAGTCTGATTCAGTTCAAGTGAATGATGACGAATATTAACGATCATTTGAG 753  
 Qy ValSerGluTyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSer 240  
 Db 754 GTCTTAAGATTAATGCTTACCAAGTTCGAAGTCACTGCGACAGCCCGCTGATTTGTTCC 813  
 Qy MetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLys-ProVa 260  
 Db 814 TTGAAGCTTAACAGTTAAATGCTCCGTCAATAGCAAGATATACATACGGGAAGCCAGT 873  
 Qy 260 LysGly-AspValThrLeuThrPheLeuProLeuSerPheTrp-GlyLysLysAsn 279  
 Db 874 GAAGGAGAGCTGTGCTTACATTAATTTTACCTTTGCTTTTGGGAGAAAGAAAGAAAT 933  
 Qy 280 IleThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluMet 299  
 Db 934 ATCAAAAAGTTTGGAGATTAATGATTTGCAAACTTCTTTTGTGATTAATTAAGAAATG 993  
 Qy 300 LysAsnValMet-----AspSerSerAsnGlyLeuSerGluTyrLeu 313  
 Db 994 AAAAAGGTAATGATTTGAAACCGCTAACGATGTCTCGAAGGAGATTAAGAAATGTG 1053  
 Qy 314 AspLeuSerSerProGlyProValGluIleLeuThrThrValThrGluSerValThrGly 333  
 Db 1054 GACCCGTCCTTTCTGGGCGCGAGAGATTAATGCTCTGAGAGATCACTCAGAGT 1113  
 Qy 334 IleSerArgAsnValSerThrAsnValPhePheLysGlnHisAspTyrIleIleGluPhe 353  
 Db 1114 ATCTCAAGATGGACAGACCAATGTGTTTTCAGACAACTGATTAATCAATTCATTAAGAT 1173  
 Qy 354 PheAspTyrThrThrValLeuLysProSerLeuAsnPheThrAlaThrValLysValThr 373  
 Db 1174 TTGATTAATAGACTGTCTTGAAACCAATCTCAACTTCACAGCCACGTGGAAGTCAAG 1233  
 Qy 374 ArgAlaAspGlyAsnGlnLeuThrLeuGluLysArgAsnAsnValIleThr-Va 393  
 Db 1234 CGCTCATGTGCAATCAACTAATCTCGAAGAAATAGAAATGATTTGGTCAAGTAC 1293  
 Qy 393 LThrGluArg-----AsnTyrThrGluTyrTrpSerGlySerAsnSerGlyAsnGln 411  
 Db 1294 GACCCAGCGGAAATTAACCACTGAG-----AGTCAGAGAGATCAAGGA 1338  
 Qy 411 SmetGluAlaValGlnLysIleAsnTyrThrValProGlnSerGlyThrPheLysIleG 431  
 Db 1339 AATGATTAATATCCAGACCGTAATTAATATCACTCCCAAAATGGAATCAATTAAGATTGA 1398  
 Qy 431 uPheProIleLeuGlnLysPheSerSerGluLeuGlnLysAlaTyrPheLeuGlySer 451  
 Db 1399 ATTCCTCAGTCATGTCCATTTCCGGCGAGCTAACCTAAGGCTTAATTTCTTGATGCTAC 1458  
 Qy 451 SserSerMetAlaValHisSerLeuPheLysSerProSerLysThrTyrIleGlnLeu 471  
 Db 1459 AAGTTCGTGACAGTCCACGATATGTTCACTCTCCAGTAAGACGTAACCACTGCTTA 1518  
 Qy 471 SThrArgAspGluAsnIleLysValGlySerProPheGluLeuValValSerGlyAsn 491  
 Db 1519 AACCAAGATGAATATATAAGGTGGGCTACCTTTGATTGATGTTGTTGTCACACAG 1578  
 Qy 491 SArgLeuLysGluLeuSerTyrMetValValSerArgGlyGlnLeuValAlaValGly 511  
 Db 1579 ACAATTCAGAGCTTAAGCTAATATGTAATCAAAAGGACAGTGTGGCTGACAGCAA 1638  
 Qy 511 ArgLysAsnSerThrMetPheSerLeuThrProGluAsnSerTrpThrProLysAlaCys 531  
 Db 1639 ACAAAAGTCAAGAACTTCTTTTAACACAGAGGCTTCCGGGCTCCAAAGGCTTAT 1698  
 Qy 531 IleValTyrTyrIleGluAspArgGlyGluIleIleSerAspValLeuLysIleProVa 551  
 Db 1699 AATGCAATTAATATTTGACAGAGATGGGAAATTAATAATCATATTTCTAAAAATCCCG 1758  
 Qy 551 LginLeuValPheLysAsnLysIleLysLeuTyrTrpSerLysValLysAlaGluProSe 571  
 Db 1759 TCAGCTTGTCTTTTGAATAAGTAAAGCTGTTTGGAGTAAACCTACAGTCAGCAAGCATC 1818

Qy 571 rGluLysValSerLeuArgIleSerValThrGlnProAspSerIleValGlyIleValAl 591  
 Db 1819 TGATTAAGTCTCTCTCAGAGATCTCTGACACAGTGTGACTCCCTGTGGGATTTGAGC 1878  
 Qy 591 aValAspLysSerValAsnLeuMetAsnAlaSerAsnAspIleThrMetGluAsnValVa 611  
 Db 1879 TGTGTGCAAAAGGTGAACATTAATGAAACCTCCACAGATTAATCAATGAAACCATGTGT 1938  
 Qy 611 HisGluLeuGlnLeuTyrAsnThrGlyTyrTyrLeuGlnLysMetPheMetAsnSerPheAl 631  
 Db 1939 CCAATGAACCTGAACCTTTATTAACACAGAAATTAATCTAGGCAATTTCACTGAACCTTCGC 1998  
 Qy 631 aValPheGlnGluCysGlyLeuTrpValLeuThrAspAlaAsnLeuThrLysAspTyrIle 651  
 Db 1999 TGATTTTCAAGAAATGTGTCTCTGGGTGTGAACATGGAACCTTATAAGATATGAT 2058  
 Qy 651 eAspGlyValTyrAspAsnAlaGluTyrAlaGluArgPheMetGluGluAsnGluLysH 671  
 Db 2059 TGATGAAGTCTACGATACGAATGAAGATTTCTGAACCTTTGACAGAGAAATGAGCCAA 2118  
 Qy 671 sIleValAspIleHisAspPheSerLeuGlySerSerProHisValArgLysHisPhePr 691  
 Db 2119 CTGTGTGATTTGAAAGATGCTTCTTCAGTTAACAAATGATACATGTCAGAGAAATTTCC 2178  
 Qy 691 oGluThrTrpIleTrpLeuAspThrAsnMetGlyTyrArgIleTyrGlnGluPheGluVa 711  
 Db 2179 AGAAACATGATTTGGCTAGACCCCTACATGGGATCTAAGATTAATGAAGAAATTTGAGT 2238  
 Qy 711 ThrValProAspSerIleThrSerTrpValAlaThrGlyPheValIleSerGluAspLe 731  
 Db 2239 TACTGTAACCTGATCCATCATCTTCTGGGTGGCTTGGTGTGATCATCAAGAACCT 2298  
 Qy 731 uGlyLeuGlyLeuThrThrProValGluLeuGlnAlaPheGlnProPheIlePhe 751  
 Db 2299 GGGTTTGGGTAAACAACGTTTCACAGACGTCGAAGCTTCAACCGTTTTCCTTT 2358  
 Qy 751 eLeuAsnLeuProTyrSerValIleArgGlyGluGluPheAlaLeuGluIleThrIlePhe 771  
 Db 2359 CTGATCTTCCCTACTCTGTTATCAAGGTGAAGATTTGCTTGAAGTATCAATCGT 2418  
 Qy 771 eAsnTyrLeuLysAspAlaThrGluValLysValIleIleGluLysSerAspLysPheAs 791  
 Db 2419 CAATTAATGAAGATACAAATTAAGCTTGATGAATCATCATGAGAAAGTCAACGCTTTGA 2478  
 Qy 791 rIleLeuMetThrSerSerGluIleAsnAlaThrGlnHisGlnGlnThrLeuLeuValPr 811  
 Db 2479 TATTTTAATGACTTCAAAATGACCAATGACACCAATATACCGGAAGACGTTCAGATTCC 2538  
 Qy 811 oSerGluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGlyGluIlePr 831  
 Db 2539 CAGGAAATAGGGGTGACTCTTGTTTTCCCAATCAAGCCACGACATTTGGGAAGATTC 2598  
 Qy 831 oIleThrValThrAlaLeuSerProThrAlaSerAspAlaValThrGlnMetIleLeuVa 851  
 Db 2599 CATCAACAGTACCGGCTGCTCAACCACTGCTGTGATGCTGTCACCAAGCAATTTGAGT 2658  
 Qy 851 LysAlaGluGlyIleGluLysSerTyrSerGlnSerIleLeuLeuAspLeuThrAspAs 871  
 Db 2659 AAAGCTGAAGGATGAAGAAAGTGAATTCGAAATGTCTTAATGATATGACAGATTAG 2718  
 Qy 871 nArgLeuGlnSerThrLeuThrLeuSerPheSerPheProProAsnThrValThrG 891  
 Db 2719 CAACGTGAAGAAAGTAACAGCAATATGAGATTTCTCTTCCCTCGATACGTCATG 2778  
 Qy 891 ySerGluArgValGlnIleThrAlaIleGluAspValLeuGlyProSerIleAsnGlyLe 911  
 Db 2779 CAGTGAAGAGTTCAGATCAACAGCAATGGAAGATATCTTGGTCTCCATCAATAGGCTT 2838  
 Qy 911 uAlaSerLeuIleArgMetProTyrGlyCysGlyGluGlnAsnMetIleAsnPheAlaPr 931  
 Db 2839 ATCTTCACGATCCGATGCTTATGATGTGTGAACAGAAATCATATATATTTGCTCC 2898

```

QY 931 cAsnIleTyrIleLeuAspTyrLeuThrIleTyrIleValGlnLeuThrAspAsnLeuLysG1 951
Db 2899 AAATATTATACATTTGGATTATCTGACTTAACAGAAACAGCTGACAGTTAATTTAAAGA 2958
QY 951 uTyAlaLeuSerPheMetArgGlnGlyTyrGlnIleArgGluLeuLeuTyrGlnArgGlnAs 971
Db 2959 AAGAGCCCTTTCATATATAGAGCAGAGTTACCAAGGAGGCTTCTCTATCAAGAGGAGAA 3018
QY 971 pGlySerPheSerAlaPheGlyAsnTyrAspProSerGlySerThrTyrLeuSerAlaPhe 991
Db 3019 TGGCTCTTTCAGTGGCTTTGGGGACATTGACTCTTCTGGAGACACTGGCTGTACAGATT 3078
QY 991 eValLeuArgCysPheLeuGlnAlaAspProTyrIleAspIleAspGlnAsnValLeuH1 1011
Db 3079 TGTTTTAAGATGCTTTCTGGAAGCTGATTAATATAGATATGATATGATAGAGATGTGTACA 3138
QY 1011 sArgThrTyrThrTyrPheLeuLysGlyIleGlnIleSerAsnGlnIlePheTyrAspProG1 1031
Db 3139 CAGAACATATACCTGGCTTAATGCACTAAGAAATTCATAGTGAATTTGGAGCCCAAG 3198
QY 1031 YArgValIleHisSerGluLeuGlnGlyValGlnIleSerProValThrLeuThrAlaTyr 1051
Db 3199 AAGAGTGAATTCACATGGAATCTCAAGGTGGCACCAAAAGCCAGTATACCTTACGCGCTTA 3258
QY 1051 rIleValThrSerLeuLeuGlyTyrArgIleTyrGlnProAsnIleAspValGlnGlnLe 1071
Db 3259 TATTTGACTTCTGCTGGAGATACAAAAGATATGATATGATATGATATGATATGATATGAT 3318
QY 1071 rIleHisPheLeuGlnSerGluPheSerArgGlyIleSerAspAsnTyrThrLeuAlaLe 1091
Db 3319 AATCAAGTTTGTGGAAATTTGAATTCAGCAGAGAAATTTGGCAATTTATACCTTGTGACAT 3378
QY 1091 uIleThrTyrAlaLeuSerSerValGlySerProIleValGlnValAlaLeuAsnMetLe 1111
Db 3379 TATATCTTATGCTCTGTCCACAGTGGGAGCCCTTAAGACAGAGAGGCTTTGAATCTGCT 3438
QY 1111 uThrTyrArgAlaGlnGlnGlnGlyIleMetGlnPheTyrValSerGlnSerLysLe 1131
Db 3439 GATTCAGCGATCAGAAAGAGAGAGAGACACAGTTCTGTTGTCATCAGACCTGCGCT 3498
QY 1131 uSerAspSerTyrGlnProArgSerLeuAspIleGlnValAlaAlaTyrAlaLeuLeuSe 1151
Db 3499 CTCTGGGCTCTGGACGCGCGCTGCGTGCATCAAAATCCGCTTATGACCTCTGCGC 3558
QY 1151 rHisPheLeuGlnPheGlnThrSerGlnGlyIleProIleMetArgTyrPheSerArgG1 1171
Db 3559 GCACACTCTGCAC---CATGTGTCAGAGGAAATCCCGGTTATGAGGTGCTCATCCAGCA 3615
QY 1171 nArgAsnSerLeuGlyGlyPheAlaSerThrGlnAspThrThrValAlaLeuLysAlaLe 1191
Db 3616 AAGAAACAGCCCTGGAGGTTTGTATCCACTCAGATATCTGTGTGGCCCTTAAGGCAATT 3675
QY 1191 uSerGluPheAlaAlaLeuMetAsnThrGlnArgThrAsnIleGlnValThrValThrG1 1211
Db 3676 ATCTGAATTTTCAGCCCTTGTGCACAGAGAAATATCAGATATCAACGACGTCACAGG 3735
QY 1211 yProSerSerPro----- 1215
Db 3736 GCCCGGCATCCCAATCCATACACTCTCCGAATGATCTCCAGAACTCTTCTCTCTCA 3795
QY 1216 ---SerProLeuAlaValAlaGlnProMetAlaValAsnIleSerAlaAsnGlyPheG1 1234
Db 3796 CCAGAGAAAGCTTCAATGCGCTAGATCCATCCATCCGTTAATGTTCTGGCCAGCGCTCGG 3855
QY 1234 yPheAlaIleCysGlnLeuAsnValValTyrAsnValIleValAlaSerGlySerArgArg 1254
Db 3856 ATTTCCTATTTCAGCTTAAGTTGACATATACGTGAAGGTTCAAGGTTCTTCTTAAGAAG 3915
QY 1254 GARArgSerIleGlnAsnGlnAlaPheAspIleAspValAlaValIleGlnLysGlnAsn 1274
Db 3916 ACGGAGATCTACCAAAATCAAGAACTTTTAATTAATGACGTCATTTGG---AATTAATCA 3972
QY 1274 sAspAspLeuAsnHisValAlaAspLeuAsnValCysThrSerPheSerGlyProGlyArgSe 1294

```

```

Db 3973 GCACGACATTTAGTACCTCGAATCTGAATGTGGCACAAGTCACTTGGCTTCAAGAGGAC 4032
QY 1294 rGlyMetAlaLeuMetGluValAlaAsnLeuLeuSerGlyPheMetValProSerGluAlaI1 1314
Db 4033 AGGCATGTCTCTTATAGAGAGTGAACCTTCTCAAGTGGCTTTAGTGCATCTTCAATTCAT 4092
QY 1314 eSerLeuSerGluThrValIleValIleValGlnIleTyrAspHisGlyIleLeuAsnLeuTyrLe 1334
Db 4093 TCTCTGTAGTAGACCCCGAAGAAAGTGAATATATATACGGAACCTCAACCTTATTTT 4152
QY 1334 uAspSerValAsnGlnIleThrGlnPheCysValAsnIleProAlaValArgAsnPhelYsVa 1354
Db 4153 AGATTCTGTAATATGATCCAGTTTGTGTAAATATCCCACTGAGAGACTACAAAGT 4212
QY 1354 lSerAsnThrGlnAspAlaSerValSerIleValAspTyrTyrGluProArgArgGlnAl 1374
Db 4213 TTCAATATTTGAGATGAGTTCATATCTGTATATGACTATATGAACCAAGAGACAGGC 4272
QY 1374 aValArgSerTyrAsnSerGluValIleLeuSerSerCysAspLeuCysSerAspValG1 1394
Db 4273 AGTGGAGAGCTTACAAACCCAGGTGAGCTGTCTCTGTTTACCTCACTCAGACACAC-- 4330
QY 1394 nGlyCysArgProCysGluAspGlyAlaSerGlySerHisHisSerSerValIlePh 1414
Db 4331 -AAGTCAAGATCACACACGAGAGGAGCCAGACTCCCTTGAGAGCTTCTTGAGAGCTTCT 4389
QY 1414 eIlePheCysPheLeuLeuTyrTyrPheMetGlu 1425
Db 4390 GGTCTTTGTCTCCCTCTCTGTACTTGTGTCAA 4423

```

## RESULT 3

```

LOCUS BC044072 4641 bp mRNA linear HTC 19-FEB-2003
DEFINITION Xenopus laevis, clone IMAGE:4681494, mRNA.
ACCESSION BC044072
VERSION BC044072.1 GI:28422351
KEYWORDS HTC.
SOURCE Xenopus laevis (African clawed frog)
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus.
1 (bases 1 to 4641)

```

## REFERENCE

```

AUTHORS Klein, S. and Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (13-JAN-2003) National Institutes of Health, Xenopus Gene
Collection (XGC), National Institute of Child Health and Human
Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD
20892-7510, USA
NIH-MGC Project

```

## REMARK

```

COMMENT Contact: XGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Igor David
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbiology.org
contact: amadan@systemsbiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha
Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAX Plate: 94 Row: D Column: 12
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, similarity but not identity to protein
This clone has the following problem: frame shifted.
Location/Qualifiers
1..4641
/organism="Xenopus laevis"
/mol_type="mRNA"

```

## FEATURES

```

source

```



```

/db xref="taxon:8355"
/clone="IMAGE:4681494"
/tissue type="Embryo, stage 31/32, Xenopus"
/clone_id="NICHD_XGC_Emb4"
/lab host="DH10B"
/notes="Vector: pCMV-SPORT6"
BASE COUNT      1322 a 1063 c 1147 g 1109 t
ORIGIN

Alignment Scores:
Pred. No.:      7e-147      Length:      4641
Score:          1480.00      Matches:      433
Percent Similarity: 46.30%      Conservative: 287
Best Local Similarity: 27.85%      Mismatches: 564
Query Match:      20.14%      Indels:      273
DB:              11      Gaps:      46

US-10-020-095-4 (1-1428) x BC044072 (1-4641)
Qy      49      LeuGluHicYbPProSerGlnValThrValLysAlaGluLeuLeuLysThrAlaSerAsn 68
      |||||      |||||      |||||      |||||      |||||      |||||
Db      226      CTGGAGGGGGGGAATGGAAAGCAAGATACAGCTCACCTACAGATGGCCAAAGTAAC 285
Qy      69      LeuThrVal-----SerValLeuGluAlaGluValPheGluLysGlySerPheLys 86
      |||||      |||||      |||||      |||||      |||||      |||||
Db      286      ACAACTGTGTAGAAAGACATCAGACAGAAATCTTCACTTCAATGTGTTCAATTGAG 345
Qy      87      ThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGluLeuTyrgluLeuArgVal 106
      |||||      |||||      |||||      |||||      |||||      |||||
Db      346      -----GTGCGGCT---CCCTCGAAGAGATAGAAAGTGGCCACCACATGCAAGT 333
Qy      107      ThrGlyArgThrGlnAspGluLeuLeuPheSerAsnSerThrArgLeuSerPheGluThr 126
      |||||      |||||      |||||      |||||      |||||      |||||
Db      394      TCATATCAGAGCGGAGAGAGACCATTA--AGCAAGACAGATTAAGTTTGTGTAAGAAA 450
Qy      127      LysArgLysSerValPheLeuGlnThrAspLysAlaLeuTyrglyLysProLysGlnGluVal 146
      |||||      |||||      |||||      |||||      |||||      |||||
Db      451      GCAGAAACCAAGTATCTTCTGTCAAACAGCAAACTTATATCAAAACAGAGACAGACAGTA 510
Qy      147      LysPheArgLysValThrLeuPheSerAspPheLysProTyrglyLysThrSerLeu--Asn 165
      |||||      |||||      |||||      |||||      |||||      |||||
Db      511      AGATTCGAGATGTTTCTTCTTAAGAAACCTCCACACAGAAAGATCAGGTGCCAAC 570
Qy      166      IleLeuLysAspProLysSerAsnLeuIleGlnGlnThrLeuSerGlnGlnSerAsp 185
      |||||      |||||      |||||      |||||      |||||      |||||
Db      571      ATGAATATACAGAGCCAGGAGAGAACCGTATGGCCAGTGGCAATGTGAGTCTCCAG 630
Qy      186      LeuGlyValLysLeuLysThrPheGlnLeuSerSerHisProIleLeuGlyAspTyrPser 205
      |||||      |||||      |||||      |||||      |||||      |||||
Db      631      CAAGGATTCGCTGACTCTCACTTCCCTCTCTCCAGCCTCCGTTGGAGAAATATTC 690
Qy      206      IleGlnValGlnValAsnAspGlnThrTyrglyLysPheGlnValSerGluTyrglyVal 225
      |||||      |||||      |||||      |||||      |||||      |||||
Db      691      ATTCAATTA-----AAGAACCTGTCCATCTTCACTGTGGAAGAAATATGTT 738
Qy      226      LeuProLysPheGluValThrLeuGlnThrProLeuTyrglySerMetLeuSerLysHis 245
      |||||      |||||      |||||      |||||      |||||      |||||
Db      739      CTGCCAGATTTAGAGTCACTCTCCACCTCCAAAAGTGTGTGTGTCACATGCGCAA 798
Qy      246      LeuAsnGlyThrIleThrAlaLysTyrglyLysProValLysGlyAspValThr 265
      |||||      |||||      |||||      |||||      |||||      |||||
Db      799      TTCCTGTATAAATCTGTGCAGGTACATTATGAAACCAAGTGCAGGGA----- 849
Qy      266      LeuThrPheLeuProLeuSerPheTyrGlyLysLysAsnIleThrLysThrPheLys 285
      |||||      |||||      |||||      |||||      |||||      |||||
Db      850      -----ACCTATTAAG 858
Qy      286      IleAsn---GlySerAlaAsnPheSerPhe-----AsnAspGluGluMet 299
      |||||      |||||      |||||      |||||      |||||      |||||
Db      859      GTGATGTGTGCGGAGCATTAACAGTACAGAAAGGCATTGGTATATGCAAAAGGAGA 918
Qy      300      LysAsnValMetAspSerSerAsnGlyLeu----- 309
      |||||      |||||      |||||      |||||      |||||      |||||

```

```

Db      919      GATCTGTGTGCTGATTTTCACTGGGAAAGCTGATCACTGTGTTGCCACATTGCATGTG 978
Qy      310      -----SerGluTyrglyLeuAspLeuSerSerProGlyProValGluLe-----LeuThr 325
      |||||      |||||      |||||      |||||      |||||      |||||
Db      979      AACTCTTCACTACATCTTCTTCAACAGCAATGGCAATGAACATTAAGGGGGGCTGCC 1038
Qy      326      ThrValThrGluSerValThrGlyLysSerArgAsnValSerThrAsnValPhePheLys 345
      |||||      |||||      |||||      |||||      |||||      |||||
Db      1039      TGGATTACAGAAAGATGGACAGGAATTCAGATTCTGCACATGTAATCCCAATTTCGA 1098
Qy      346      GlnHisAspTyrglyLeuGluPhePheAspTyrglyThrValLeuLysProSerLeuAsn 365
      |||||      |||||      |||||      |||||      |||||      |||||
Db      1099      TATGACCTCAACAGTGTCACTTCTCGAATGCTGATGATTAATTAAGCTGGGATTCGA 1158
Qy      366      PheThrAlaThrValLysValThrArgAlaAspGlyAsnGlnLeuThr----- 381
      |||||      |||||      |||||      |||||      |||||      |||||
Db      1159      TACATCGGAGACAGTGAAGTTGTAGATGCCAGTGAACCCCATCCCAAAATCCAGTG 1218
Qy      382      --LeuGluLysArgAsnAsnValValIleThr--ValThrGlnArgAsnTyrglyThr 399
      |||||      |||||      |||||      |||||      |||||      |||||
Db      1219      TATCTGATGCACACGATGGAATAGTAACATGACCTTGATGACCATTAAGATGACATT 1278
Qy      400      GluTyrglyTyrPserGlySerAsnSerGlyAsnGlnLysMetGluAlaValGlnLysIleAsn 419
      |||||      |||||      |||||      |||||      |||||      |||||
Db      1279      GCTCCCTTCAAGCTGACATACAGACAGATTGGAAAGCA----- 1317
Qy      420      TyrglyValProGlnSerGlyThrPheLysIleGluPheProIleLeuGlu----- 436
      |||||      |||||      |||||      |||||      |||||      |||||
Db      1318      ---ACGGTGTCCCTGACCGGCCAAAGACAGTGAAGAGACCTTTATGCAATGACTAC 1374
Qy      437      -----AspSerSerGluLeuGlnLeuLysAlaTyrglyPheLeuGly 449
      |||||      |||||      |||||      |||||      |||||      |||||
Db      1375      GCGATGATGTCCTTCAAGCATGGGTAGCGATGCTTAAGCTCAACCGCTTTATATCCCG 1434
Qy      450      SerLysSerSerMetAlaValHisSerLeu-----PheLysSerProSerLysThr 466
      |||||      |||||      |||||      |||||      |||||      |||||
Db      1435      AGCAAGAGTTTATTAAGCTCCACCTCCAGAAATGGGAGTTCTGTGTAAGGGCAACAA 1494
Qy      467      TyrglyLeuLysThrArgAspGluAsnIleLysValGlySerProPheGluLeuVal 486
      |||||      |||||      |||||      |||||      |||||      |||||
Db      1495      GAAGTCCAGGTGAATATCTTATCAAGCCACAGACTTGGAGAG----- 1539
Qy      487      ValSerGlyAsnLysArgLeuLysGluLeuSerTyrglyMetValValSerArgGlyGlnLeu 506
      |||||      |||||      |||||      |||||      |||||      |||||
Db      1540      ---GAGCAAAAGACATTGAGTACACTAGCTGGCAGTTCTTAAGGTGCCATA 1590
Qy      507      -----ValAlaValGlyLysGlnAsnSerThrMet----- 516
      |||||      |||||      |||||      |||||      |||||      |||||
Db      1591      AAGAAATCTGACGTTTAAAGATTCATTAATTAATGATGTGAAGTCTTGAGAGCGGAG 1650
Qy      517      -----PheSerLeuThrProGluAsnSerTyrglyProLysAlaCysValIleValTyrgly 534
      |||||      |||||      |||||      |||||      |||||      |||||
Db      1651      GTTCCATCAAACTCAAGCTCAAGTGTGACGTCTCTCCACACTTCTGCAATCTTGACCTTAC 1710
Qy      535      TyrglyLeuAspAspGlyGluLeuLysSerAspValLeuLysIleProValGlnLeuVal 554
      |||||      |||||      |||||      |||||      |||||      |||||
Db      1711      GTTTATCTCCCGTGTGAGAGATGTGTGCAAGATGCAAAAGTTCAAGCTACAAAGTGC 1770
Qy      555      PheLysAsnLysIleLysLeuTyrglyPserLysValLysAlaGluProSerGluLysVal 574
      |||||      |||||      |||||      |||||      |||||      |||||
Db      1771      TTTAGAAATTAAGGTGACATGTGCTTCTCCAGATGAAGTCTTACACAGATCAAGTGT 1830
Qy      575      SerLeuAsnIleSerValThrGlnProAspSerIleValGlyIleValAlaValAspLys 594
      |||||      |||||      |||||      |||||      |||||      |||||
Db      1831      TCCCTCAAGTTCAGGCACT---GCCGGGTCTCTGTGCTGCTGAGAGTGTGTGATTAAG 1887
Qy      595      SerValAsnLysMetAsnAlaSerAsnAspIleThrMetGluAsnVal----- 610
      |||||      |||||      |||||      |||||      |||||      |||||
Db      1888      AGTGTGTGTATATGAACCTGACCAAGAACTGACCGCGCAAGAAATTCGGGAGCTCTTC 1947
Qy      611      -----ValHisGluLeuGlu----- 615
      |||||      |||||      |||||      |||||      |||||      |||||
Db      1948      CCAGTCAATGATTTTGGAAATTATGATTATCGGATTCAGAAAGCTGAGATCATCTGTCCA 2007

```

QY 616 -----LeuTyraAsnThr 619  
 Db 2008 TTCGACCCCTTCTTCCTGGACCATTCATAGAAATTTTCTAGATTCATATATATATCC 2067  
 QY 620 GlyTyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTrp 639  
 Db 2068 CCATGGCATATGCCAGAGAACGTGATGTCTACGGCTCTCTCAAGAAAGTGCACCTAA 2127  
 QY 640 ValLeuThrAspAlaAsnLeuThrTyrAspTyrIleAspGlyValTyrAsp----- 656  
 Db 2128 ATCATCACCAGGCTGCACCTTAAAGAACAGTGTGTGCAACATTAATGATTTTAAAG 2187  
 QY 657 -----AsnAlaGluTyrAlaGluArgPheMetGluGlnAsnGluYHisIleValAsp 674  
 Db 2188 GCCAGGGCTACAGAGCTTCAGAGCATCTGAGACCGAATGATTTG-CATTGTCTCC 2246  
 QY 675 IleHisAspPheSerLeuGlySerSerProHis----- 685  
 Db 2247 CTGGGATA-TCGTGCGCGGATCTATACCAACAGCTACCAAGACAGAGCTGCACGC 2305  
 QY 686 -----ValArgLysHisPheProGluThrTrpIleTrpLeuAspThrAsnMetGly 702  
 Db 2306 GTTGAATATATCCGAACCTACTTTCAGAGAGCTGATATGGAGCTGGCATCTGTGGGC 2365  
 QY 703 TyrArgIleTyrGlnGluPheGluValThrValProAspSerIleThrSerTrpValAla 722  
 Db 2366 GAGTGTGGTACAGAGAGCTCATCGCGAGTCCAGACACATCAGACACTGGAATGCT 2425  
 QY 723 ThrGlyPheValIleSerGluAspLeuGlyLeuThrThrProValGluLeu 742  
 Db 2426 GAGAGCTTCTGTATGGAGCCAGT---GGATTGGGATC---TCTTCAACCACTCCCT 2479  
 QY 743 GlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArgGlyGlu 762  
 Db 2480 CGAGCTTCACGCCCTCTTGTGTGAGAGCTCACTGCGCTGTGTGTGGAGAGAG 2539  
 QY 763 GluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrGluValIleVal 782  
 Db 2540 TCATTGCTCTCAAGGCACTGTCTTAACTACCTGAAACAGCTGATTAAGTTACGACA 2599  
 QY 783 IleIleGluLysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsnAlaThr 802  
 Db 2600 GCCCTGAGAGCTTCCAGAGAGTTGGAG-----GAGGAACCATGTGCT 2641  
 QY 803 GlyHisGlnGlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPheProIle 822  
 Db 2642 GACTGTCAGTATAGAGCTGTGTGTGTGACAGAGAGCAAGACCTTCTACAGAACGT 2701  
 QY 823 ArgProThrHisLeuGlyGluIleProIleThrValIleLeuSer----- 838  
 Db 2702 AAAGCTTCACTGAGAGAGGTGAATGTCACGGTAAAGACGGTGTCTGATCTCAG 2761  
 QY 839 -----ProThrAlaSerAspAlaValThr 846  
 Db 2762 GACTTGCAATATAGAAATCCCTATAGTCCCTAAACAGGGAGATCAGACCTGAT 2821  
 QY 847 GlnMetIleLeuValIleAlaGluGlyLe-----GluLysSerTyrSerGlnSerIle 864  
 Db 2822 AAGCCACTACTGTGTCAGCAGGAGGTCTCTGAAAGAAACCTCAGACCTTCTGCTC 2881  
 QY 865 LeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeuSerPheSerPhe 884  
 Db 2882 TGCTGCCAAGCTGTGATGATCATCTTAAACA-----GAGAGATATCTCTTAAGTC 2935  
 QY 885 ProProAsnThrValThrLysSerGluArgValGlnIleThrAlaIleGlyAspValLeu 904  
 Db 2936 CTTGAATATATCTCAAGGATTCAGAAAGACATATGACGTCTGAGATCTGATG 2995  
 QY 905 GlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGlyCysGlyGluGln 924  
 Db 2996 GGAACAGCATGACAGAACTGAGCCGTCTCTGCGCATGCCGTATGATGTGGAGAGAG 3055

QY 925 AsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThrLysLysGln 944  
 Db 3056 AACATGTTCTCTTGGCCCCAATATATCTTCAATCTTCAAGTATCTGAAAGAACATCACC 3115  
 QY 945 LeuThrAspAsnLeuLysGlyValAlaLeuSerPheMetArgGlnGlyTyrGlnArgGlu 964  
 Db 3116 CTGAGCCCTGATGCCAGAGCAAGCCAAAGGTTCCTTGAAGGGGCTCCAGAGCAG 3175  
 QY 965 LeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyrAspProSerGly 984  
 Db 3176 CTGACATCAAAACGGATGACGATCATACAGCCGATTTGGGCAAGAGCAAGAGGGG 3235  
 QY 985 SerThrTrpLeuSerAlaPheValLeuArgCysPheLeuGlnAlaAspProTyrIleAsp 1004  
 Db 3236 AACACATGCTGACGTGATTTGTGTGAAGTCTTCAAGCAAGGCTCGTCCCTACATCTTC 3295  
 QY 1005 IleAspGlnAsnValLeuHisArgThrTyrThrTrpLeuLysGlyHisGlnLysSerAsn 1024  
 Db 3296 ATAGATGAAGTCACTTAAACATTTCTTCACTTGGCTGAAGAAACATCAACATGAAGT 3355  
 QY 1025 GlyLysPheTrpAspProGlyArgValIleHisSerGluLeuGlnGlyValAsnLysSer 1044  
 Db 3356 GCGTGTTCAGAGTGTGGGAGAGCTTTCATATATGCTATGAAGGAGGAAATGATGAT 3415  
 QY 1045 ProValThrLeuThrAlaTyrIleValThrSerLeuGlnGlyTyrArgLysTyrGlnPro 1064  
 Db 3416 GAACCTCTCTCTGCTTATGTAACAATGCTTTGCTA-----GAACCT 3460  
 QY 1065 AsnIleAspValGlnGluSerIleHisPheLeuGlnSerGluPheSer----- 1080  
 Db 3461 GGTCTATCTGTACAGAGCCCAATG-----GTGAGGGTGTCTTATCTGCTGAGAGAG 3514  
 QY 1081 -----ArgGlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSer 1098  
 Db 3515 GCTGCCAAAGATGTAGAGATATCTTACCCAGAGCTGTGCTGAGCTTACCTTCACTCTG 3574  
 QY 1099 ValGlySerProLysAlaLysGlu---AlaLeuAsnMetLeuThrTrpArgAlaGluGln 1117  
 Db 3575 AGTGTGATATCTGAACCTCAGGAGATGCTGTGCTGAACCTGATGAGAGAGCTGTAA 3634  
 QY 1118 GluGlyGlyMetGlnPheTrp-----ValSerSerGluSerLysSerAsp 1133  
 Db 3635 AGTGAAGAGACAGTTACCTGGAAAGAAACCTGCGGAGCTCCTCTGATTAACCTTAC 3694  
 QY 1134 SerTrpGlnProArgSerLeuAspIleGluValAlaAlaTyrAlaLeuLeuSerHisPhe 1153  
 Db 3695 TGTATCCAGGCTCCCTCAGCGAGGTGAGTTAACCTTCTATGTACTGTTAACCTTCTC 3754  
 QY 1154 -----LeuGlnPheGlnThrSerGluGlyIleProIleMetArgTrpLeuSerArgGln 1171  
 Db 3755 TCTGGCCTCAAAAAGAGCTGGGAGAACTTCTGATGATGATCAACTGCTGACGACAG 3814  
 QY 1172 ArgAsnSerLeuGlyGlyPheAlaSerThrGlnAspThrThrValAlaLeuLysAlaLeu 1191  
 Db 3815 CAATATCATATGGGGATCTCTTACTACAGACAGGGTGTAGCTCTACAGGCTCTT 3874  
 QY 1192 SerGluPheAlaAlaLeuMetAsnThrGluArgThrAsnIleGlnValThrValThrGly 1211  
 Db 3875 GCTAAGTATGCCAGGCCACATTCAGCAAGGAGATGTAAACATCAGTCAAGTCTCC 3934  
 QY 1212 ProSer-----SerPro 1215  
 Db 3935 AAGACTGATTCATCAACAGTTCATATGATGATCAAAATACCGGCTCTCTCTACAGAA 3994  
 QY 1216 SerProLeuAlaValAlaGlnProMetAlaValAsnIleSerAlaAsnGlyPheGlyPhe 1235  
 Db 3995 TCTTCACTGTCAAGATATC---CTGGGACATTACTCTGTGTACAGCAAGGAGGAGGCTGT 4051  
 QY 1236 AlaIleCysGlnLeuAsnValValTyrAsnValLysAlaSerGlySerSerArgArgArg 1255  
 Db 4052 GTGATATGACAGAGTGTCTGATATCAACATTCCTCCACCT----- 4093  
 QY 1256 ArgSerIleGlnGlnGluAlaPheAspLeuAspValAlaValLysGluAsnLys--- 1274

```

Db      4094 -----AGAACTGATGCCACCTTCTCACTATCAAGTGAAGACACTACCAACAATGT 4144
Qy      1275 -----AspAerLeuAsnHisValAspLeuValCysThrSerPheSerGlyPro--- 1231
Db      4145 CCTCAAGATCTATTAAATTCATGACCTTACGACTCTGTAATATACAGGAACCTCG 4204
Qy      1292 G1yArSerG1yMetAlaLeuMetG1yValAsnLeuLeuSerG1yPheMetValProSer 1311
Db      4205 GAGAATCAACATGCTGTGATTAAGTGAAGTGTCTCTGATTT---ATCCCATCA 4261
Qy      1312 G1yAla-----IleSerLeuSerG1yThrValLysLysValG1yThrAspHis 1327
Db      4262 AAGGCTCAATCAGATCTGAGAGAGCAACAACTAAAGCCGAGTACATTCAGACA 4321
Qy      1328 G1yLysLeuAsnLeuTyrLeuAspSerValAsnG1yThrG1yPheCysValAsnIlePro 1347
Db      4322 GACATGTCGATTAATTAATTTGATGAGCTGGACATGACCTCTCCATCTGCCTTACG 4381
Qy      1348 AlaValArgAsnPhenylValSerAsnThrG1yAspAlaSerValSerIleValAspTyr 1367
Db      4382 GTGACACAGACATTTGATGAGAGACCTAAAGCCGACAGTGAAGTATGATTTAT 4441
Qy      1368 TyrG1yProAlaG1yG1yAlaValArgSerTyrAsnSerG1yValLysLeuSerSerCys 1387
Db      4442 TATGAGACAGAGAGAGGAGGAGTTCGCGAGTACAACTCC----- 4480
Qy      1388 AspLeuCysSerAspValG1yG1yCysArgProCys---G1yAspG1yAlaSerGly--- 1405
Db      4481 -----CCGTCAGCTCTGATGGAGAGCAAGGAC 4510
Qy      1406 -----SerHis 1408
Db      4511 ACAAAATGAGATGTCTCCGCGACATTAACAGAACTAGTGAAGCAATCATCAAC 4570
Qy      1409 HisSerValIlePheIlePheCysPheLysLeuTyrPhe 1423
Db      4571 AGATCTGCTCAATATATTTTACATATATGAATTAATATATGTTT 4615

RESULT 4
BX391819/c 941 bp mRNA linear EST 13-MAY-2003
LOCUS BX391819 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA
DEFINITION BX391819 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA
ACCESSION BX391819
VERSION BX391819.1 GI:30627404
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 941)
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 3578.f for
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0BA1029ZD07_CS02740_1&cluster=3578.f.
Contact : Feng Liang Email : fliang@life-tech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0BA1029ZD07_CS02740_1.
FEATURES
SOURCE
1..941
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1081YB18"

```

```

/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_id="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 271 a 214 c 181 g 266 t 9 others
ORIGIN
Alignment Scores:
Pred. No.: 6 01e-146 Length: 941
Score: 1461.00 Matches: 293
Percent Similarity: 93.95% Conservative: 2
Best Local Similarity: 93.31% Mismatches: 17
Query Match: 19.88% Indels: 3
DB: Gaps: 0
US-10-020-095-4 (1-1428) x BX391819 (1-941)
Qy      941 LysLysLysG1yLeuThrAspAsnLeuLysG1yValAlaLeuSerPheMetArgG1y 960
Db      938 AAAAGNANCANCTGACAGATATTTGAAAAGAAAGCTGTTTCATTATGAGGCAAGGG 879
Qy      961 TyrG1yArgG1yLeuLeuTyrG1yArgG1yAspGlySerPheSerAlaPheG1yAsnTyr 980
Db      878 TACCAGAGACACTTCTCTATCAGAGGAGATGGCTC-TTCACTGCTTTTGGAAATTA 820
Qy      981 AspProSerGlySerThrTrpLeuSerAlaPheValLeuArgCysPheLeuG1yAlaAsp 1000
Db      819 GACCTCTGGAGCACTGTGTGTGCTGCTTTGTTTAAGATGTTCTTGAAAGCCGAT 760
Qy      1001 ProTyrIleAspIleAspG1yAsnValLeuHisArgThrTyrThrTrpLeuLysG1yHis 1020
Db      759 CTTTACATATATATTTGATGAGATGTGTTACAGAACTATCACTGCTTAAAGACAT 700
Qy      1021 G1yLysSerAsnG1yG1yPheTrpAspProG1yArgValIleHisSerG1yLeuG1y 1040
Db      699 CAGAAATCCAAACGTTGATTTGGATCCAGAGAGTGAATCATATGATGAGCTTCAAGT 640
Qy      1041 G1yAsnLysSerProValThrIleThrAlaTyrIleValHisSerLeuLeuG1yTyrArg 1060
Db      639 GGCATTAAGATCCAGTACACTTACAGCTTAATATGTAATCTCTCTCGGATATAGA 580
Qy      1061 LysTyrG1yProAsnIleAspValG1yG1ySerIleHisPheLeuG1ySerG1yPheSer 1080
Db      579 AAGTATCAGCTTAATCATGATGAGCAAGTCTATCATTTTGGAGTGTGAATCAGT 520
Qy      1081 ArgG1yLysSerAspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGly 1100
Db      519 AGAGGAATTCACAGCAATTAATCTAGCCCTTATTAATTAATCATATGTCATCAGTGGG 460
Qy      1101 SerProLysAlaLysG1yAlaLeuAsnMetLeuThrTrpArgAlaG1yG1yG1y 1120
Db      459 AGTCTTAAGCCGAGAGAGTGTGAATATGCTACTTGAGACAGCAAGAGG1GGC 400
Qy      1121 MetG1yPheTrpValSerSerG1ySerLysLeuSerAspSerTrpG1yProAlaSerLeu 1140
Db      399 ATGCATATTCGGGTGTCATAGAGTCCAAACTTTCGATCCCTGGACGACGCTCCCTG 340
Qy      1141 AspIleG1yValAlaLysTyrAlaLeuLeuSerHisPheLeuG1yPheG1yThrSerG1y 1160
Db      339 GATATGAGATTGACACCTGATGACATGCTCTCACATCTTCAATTCAGACTTCTGAG 280
Qy      1161 G1yTyrIleProIleMetArgTrpLeuSerArgG1yArgAsnSerLeuG1yG1yPheAlaSer 1180
Db      279 GGAATCCCAATTTATGAGTGGCTTAACAGGCAAGAAATAGCTTGAGTGTGCTTTCATCT 220
Qy      1181 ThrG1yAspThrThrValAlaLeuLysValAlaLeuSerG1yPheAlaAlaLeuMetAsnThr 1200
Db      219 ACTCAGATACCACTGTGCTTTAAAGCTCTGTCTGAATTTGACGCTTAAGATACA 160
Qy      1201 G1yArgThrAsnIleG1yValThrValThrG1yProSerSerProSerProLeuAlaVal 1220

```

Db	159	GAAAGGACAAATATCATCAAGTACCGTGACGGGGCCCTAGCTACCAAGTCTCTTGCTGTG	100
Oy	1221	ValGlnPFromEalValAsn1IleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeu	1244
Db	99	GTACAGCCCAACGGCGCATTAATATTTCCGCAAAATGGTTTGATTGCTATTGTTCAGCTC	40
Oy	1241	Asn--ValValTyrAsnVallyAsAlaSerGlySerSer	1252
Db	39	AATGNTTGTATATATATGATGAGCGCTTCTGGGCTCTTCT	2
RESULT 5			
AL554092		1201 bp	mRNA
LOCUS			linear
DEFINITION	AL554092 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA		EST 31-MAY-2001
ACCESSION	AL554092		
VERSION	AL554092.2		
KEYWORDS	GI:31275905		
SOURCE	EST.		
ORGANISM	Homo sapiens (human)		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
JOURNAL	1 (bases 1 to 1201)		
COMMENT	Li, W.B., Gruber, C., Jessee, J. and Polayes, D. Full-length cDNA libraries and normalization Unpublished		
COMMENT	On Feb 15, 2001 this sequence version replaced gi:12894544.		
COMMENT	Contact: Genoscope		
COMMENT	Genoscope - Centre National de Sequencage		
COMMENT	BP 191 91006 EVRY cedex - France		
COMMENT	Email: seqref@genoscope.cns.fr Web : www.genoscope.cns.fr		
COMMENT	Library was constructed by Life Technologies, a division of		
COMMENT	Invitrogen. This sequence belongs to sequence cluster 3578.f For		
COMMENT	more information about this cluster, see		
COMMENT	http://www.genoscope.cns.fr/		
COMMENT	cgi-bin/cluster.cgi?seq=CSODI081DA090P1c&cluster=3578.f. Contact :		
COMMENT	Reng Liang Email : lliang@life-tech.com URL :		
COMMENT	http://fulllength.invitrogen.com/ Invitrogen Corporation 1600		
COMMENT	Faraday Avenue Genoscope sequence ID : CSODI081DA090P1.		
FEATURES			
source			
	1..1201		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="CSODI081YB18"		
	/tissue_type="PLACENTA COT 25-NORMALIZED"		
	/clone_id="Homo sapiens PLACENTA COT 25-NORMALIZED"		
	/note="1st strand cDNA was primed with a NotI-oligo(dT)		
	primer. Five prime end enriched. Double-strand cDNA was		
	digested with Not I and cloned into the Not I and EcoR V		
	sites of the pCMVSPORT 6 vector. Library was normalized."		
	BASE COUNT		
	359 a 238 c 236 g 293 t		
	75 others		
ORIGIN			
Alignment Scores:			
Pred. No.:	7.4e-140	Length:	1201
Score:	1406.00	Matches:	293
Percent Similarity:	90.74%	Conservative:	1
Best Local Similarity:	90.43%	Mismatches:	29
Query Match:	19.13%	Indels:	4
Db:	9	Gaps:	1
US-10-020-095-4 (1-1428) x AL554092 (1-1201)			
Oy	1	MetGlnGlyProProLeuLeuThrAlaAlaHisLeuLeuGlyValCysThrAlaAlaLeu	20
Db	92	ATGACGGGACCCACCGCTCTGTCAMCGCMGCCACCTCTGCTGTGTGACACCGCCGCGTG	151
Oy	21	AlaValAlaProGlyProArgPheLeuValThrAlaProGlyValIleLeuArgProGlyGly	40
Db	152	GCCGGGCTCCCGGGCTCTGGTTTCTGTGTGAMAACACAGGGATATATCAGGCGCCGAGGA	211
Oy	41	AsnValThrIleGlyValGlnLeuLeuGlnHisCysProSerGlnValThrValIlyAla	60

Db	212	ATGTGACTATTGGAGTGGAGCTTCTGGAATACCTGCTTAAAGTACTGTGAAGCG	271
Qy	61	GlutLeuLeuYsThrAlaSerAsnLeuThrValSerValLeuGluAlaGluYValPhe	80
Db	272	GAGCTGCTCAAGACAGCATCAAAAMTCACTCTCTCTGTGTCGGAAGCAAGAGACTCTT	331
Qy	81	GluYsGlySerPheYsThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGlu	100
Db	332	GAATAAGCTCTTTTAARAAACTTAMTCTTCATTAATACMTCTGAAACATGCAATGAG	391
Qy	101	IleYrGluLeuArgValThrGlyArgThrGlnAspGluIleuPheSerAsnSerThr	120
Db	392	ATTATATGCTATGCTATGTAACCGACAGTACCAAGATGTGAATTTATTTCTTAAMGTACC	451
Qy	121	ArgLeuSerPheGluThrTrpYsArgIleSerValPheIleGlnThrAspGlyAlaLeuYr	140
Db	452	CGCTTATCATTTTGAGACCAAGAGATATCTGTTTTCATTCAACAGAAAGAGCCTTATAC	511
Qy	141	LysProLysGlnGluValLysPheArgIleValThrLeuPheSerAspPheLysProYr	160
Db	512	AAGCAAAAGCAAGAGTGAAGTTTCGCAATGTTCACCTTCTCAGATTTTAAAGCTTAC	571
Qy	161	LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGlnThrLeu	180
Db	572	AAAACCTCTTAAACATTTCTCATTAAGAGACCCCAAAATCAATTTGATCAACAGAGTGTG	631
Qy	181	SerGlnGlnSerAspLeuGlyValIleSerYsThrPheGlnLeuSerSerHisProIle	200
Db	632	TCACACAAAGATGATCTTGGAGTCACTTTCCAAACTTTTCAGCTTCTTCCATTCATTA	691
Qy	201	LeuGlyAspTrpSerIleGlnValGlnValAsnAspGlnThrYrYrGlnSerPheGln	220
Db	692	CTTGATGACTGTGTCTATTCAATTCAGTCAAGTGAATGACGACATATTCATTCATTTGAG	751
Qy	221	ValSerGluYrValLeuProLysPheGluValThrLeuGlnThrProLeuYrCySer	240
Db	752	GTTTCAGAAATGATTAACCAAAATTTGAAAGTCACTTGCAGACACCACTTATATGTTCT	811
Qy	241	MetAsnSerLysHisLeuAsnGlyThrIleThrAlaYsYrThrYrGlyLysProVal	260
Db	812	ATGATATCTTAAGCATTTTAATATGTAATCAAC-ACGGCAAGATATCAATATGGAGACCACTG	870
Qy	261	LysGlyAspValThrLeuThrPheLeuProLeuSerPheTrpGlyLysAlaSerAsnIle	280
Db	871	AAAGAGACGTAACGCTTACATTTTACCTTATCCTTTTGGGGAAGAAAGAAATATTT	930
Qy	281	ThrYsThrPheYsIleAsnGlySerAlaAsnPheSerPheAsnAspGluIleLys	300
Db	931	ACAAAAACATTTAAGATTAATATGATCTCGCAAACTTCTTTTAAATGAAAGAGATGTAA	990
Qy	301	AsnValMetAspSerSerAsnGlyLeuSerGluYrLeuAspLeuSerSerProGlyPro	320
Db	991	AM-GTAAATGATCT-TCAAATGACCTTTCTGAA--ACYKGAATATATTTCCCTCGACM	1045
Qy	321	ValGluIleLeu 324	
Db	1046	GTAGAAATTTTA 1057	
RESULT 6			
BC049868			
LOCUS	BC049868	4667 bp	mRNA
DEFINITION	Mus musculus, murinoglobulin 1, clone IMAGE:5097177, mRNA.		
ACCESSION	BC049868		
VERSION	BC049868.1	GI:29437204	
KEYWORDS	HTC.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Euteheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
TITLE	1 (bases 1 to 4667)		
	Strauberg,R.		
	Direct Submission		

## JOURNAL

Submitted (31-MAR-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK  
COMMENT  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: Jeffrey E. Green, M.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)  
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),  
Gaithersburg, Maryland;  
Web site: <http://www.nisc.nih.gov/>  
Contact: [nisc\\_mgc@nih.gov](mailto:nisc_mgc@nih.gov)  
Ahkerley, N., Ayala, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakeley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaapi, R., Madu, O.L., Masfello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantirpop, S., Thomas, P.J., Touchman, J.W., Turgerson, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LIML at: <http://image.liml.gov>  
Series: IRAC Plate: 99 Row: h Column: 5  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6678965  
This clone has the following problem: frame shifted.

## FEATURES

## SOURCE

1. 4667  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5097177"  
/tissue\_type="Liver, normal, 5 month old male mouse."  
/clone\_id="NCI CGAP L49"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"  
BASE COUNT 1346 a 1126 c 1046 g 1149 t  
ORIGIN

## Alignment Scores:

Pred. No.: 7.39e-139 Length: 4667  
Score: 1405.50 Matches: 445  
Percent Similarity: 45.10% Conservative: 268  
Best Local Similarity: 28.15% Mismatches: 559  
Query Match: 19.13% Indels: 310  
DB: 11 Gaps: 53

US-10-020-095-4 (1-1428) x BC049868 (1-4667)

QY 4 ProProleuLeuThraAlaAlaLeu-----LeuCyValCys 16  
Db 73 CTTCTGCTCTTCTTACACACATGTGGAAGACAGAGGGCTCAGTTGTGCTTTT 132  
QY 17 ThraAlaAlaLeuA-----ValAlaProGlyProAlaGpPheLeu 29  
Db 133 TCGGTTCTCTCTGCTTCGTCATCTGCTCTTTACTCAACGAGACTCGAAGTATATG 192  
QY 30 ValThraAlaProGlyLeileArgProGlyGlyAsnValThrlLeglValAlaGluLeu 49  
Db 193 GTGCTGTCTCTT-----CACTTAC 216  
QY 50 GluHisCyProSerGlnValThraValysAlaGluLeuLeuYsrThraAsnLeu 69  
Db 217 ACTGAGACCCCTGGAATAATCTGCTCATCTATACAGCTGAATGAACAGTACCGTC 276  
QY 70 ThraValserValleuGluAlaGluGly-----ValPheGlu 81  
Db 277 ACAGCTTCCTGTATCTCAATCAGAAAGAAACCTGTTGCATGAGCTGTGCTTAC 336

QY 82 LysGlySerPheLeuThrlLeuThr-----LeuProSerLeuProLeuAsnSerAlaAsp 99  
Db 337 AAGACCTGTTCACAAAGTGTTCCTTCATATTCCTAGACTCTCCCTTCTGTATGAGAG 396  
QY 100 GluIleThrlGluLeuArgValThrlGlyArgThrlGlnAspGluIleLeuPheSerAsnSer 119  
Db 397 GATTTCTCTACCTGACATCAAAAGGGCCACCCAGAA-----TTCAGCAAAAGG 447  
QY 120 ThraGluSerPheGluThrlLysArgIleSerValPheIleGlnThraAspValAlaLeu 139  
Db 448 AAGCAGCTGCTGTGAAGAACAAAGAAAGTGTGTCTTGTGTGACACATTAAGCCCGTC 507  
QY 140 TrrLysProLysGlnGluValLysPheArgIleValThrlLeuPheSerAspPheLysPro 159  
Db 508 TACAAACCAAGCAATCAGTTAAATTCGGGTTGTCTATGATTAATATGATCACTGCC 567  
QY 160 TrrLysThrlSerLeuAsnLe-----LeuIleLysAspProLysSerAsnLeuIleGln 178  
Db 568 CTGAATGAGTTGCTTCTGCTTACATTCAGGATCCAAAGAAAGCAATATATGACAG 627  
QY 179 TrpLeuSerGlnIleAsnAspLeuGlyValIleSerLysThrlPheGlnLeuSerSerHis 198  
Db 628 TGAGGAGATTTAAGCTGAGAAATGGGCTTAAGCAAAATGCTTCAGCTGGACAGAG 687  
QY 199 ProIleLeuGlyAspTrpSerIleGlnVal-----GlnValAsnAspGlnThrlTrrGln 217  
Db 688 CCATTCAGAGGCCCTTACAAAGATAGTGTGCACAAAGATCAGGGAGAAAGAAACAC 747  
QY 218 SerPheGlnValSerGluTrrValLeuProLysPheGluValThrlLeuGlnThrlProLeu 237  
Db 748 TCTTTTACGTGATGAAATTTGTGCTCCAGATTAAATGTCGACCTGAAGCCCAAT 807  
QY 238 TrrCySerSerMetAsnSerLysHisLeuAsnGlyThrlLeThraLysTrrThrlGly 257  
Db 808 GCCATGCTGTGATATGATGAGTCTCAGTGTGACATGTCGGAATATACCTATGGG 867  
QY 258 LysProValLysGlyAspValThrlLeuThrlPheLeuProLeuSerPheTrpGlyLys 277  
Db 868 AACCTGTTCAGACATGTG----- 888  
QY 278 LysAsnIleThrlLysThrlPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGlu 297  
Db 889 -----AAGATMAACGTCTGCGGTGAGACTGAGCTGATGCAGA 927  
QY 298 GluMetLysAsnValMetAspSerSerAsnGlyLeuSer-----GluTrrLeu 313  
Db 928 GAAGTAAATTTCTAGCTAGAC--AACATGGCTGACATACACAAAGTGAACATCACT 984  
QY 314 AspLeuSerSerProGlyPro-----ValGluIleLeuThrlVal 327  
Db 985 GAGCTCCATCAAAAGAAAGAAAGAAATATGAGCTTCAGCTTTTCCATGTGAATGCATGTT 1044  
QY 328 ThrlGluSerValThrlGlyLe-----SerArgAsnValSerThraAsnValPhePheLys 345  
Db 1045 ACAGAGAAAGGACAGAGATTTGAGTTCACTGCTGGAACAACCTAAATTTGAAGAAATC 1104  
QY 346 GlnHisAspTrrIleIleGluPhePheAspTrrThrlValLeuLysProSerLeuAsn 365  
Db 1105 ACAAAAGACTCTA-----TTCTGAAGGACAGATTCACACTTCACACATGGAATTCGA 1158  
QY 366 PheThraLysThraValLysValThraArgAlaAspGlyAsnGlnIleuThrlLeuGluArg 385  
Db 1159 TTCTTTGTGAAGATCCGCTAGTGAATATCAAGAGACCCATATCCCAATAGAGAAA--- 1215  
QY 386 ArgAsnAsnValAlleThraValThrlGlnAsnValTrrGluTrrSerGlySer 405  
Db 1216 -----GTCTTCATCAAGACAAAGAACTTAGCTATACAGTCTACTACCACTGAT 1266  
QY 406 AsnSerGlyAsnGlnLysMetGluAlaValAlaGlnLysIleAsnTrrThraValProGlnSer 425  
Db 1267 CAGCATGGC-----CTG 1278

QY 426 GlyThrPheLysIleGluPheProIleLeuGluAspSerSerGluLeuGlnLysAla 445  
 DB 1279 GCAGAGTCTCCATGATGATCCACATGATCTCGGGCTCTTCC---CTCCATTTCAAAGTC 1335  
 QY 446 -----TyrPheGluGlySerLysSerMet 454  
 DB 1336 AACCCAAAGAGGAAGATTCATGTTCTTATTTCTATGTCATGAGAAAGACATCGAAGT 1395  
 QY 455 AlaValHisSerLeuPheLys-----SerProSerLysThrTyrIleGlnLeuLysThr 472  
 DB 1396 GGAAGACATGTCGCTATGCTGTTTACTCCACAGAAAGAGCTATCACTTACTTACACACA 1455  
 QY 473 ArgAspGluAsnIle-----LysValGlySerProPheGluLeu 485  
 DB 1456 GAGACAGACGATCTTGCCCTGCAACACGATTCACACAGTTCAGGACATTTT----- 1509  
 QY 486 ValValSerGlyAsnLysArg---LeuLysGluLeu-----SerTyrMetAlaValSer 502  
 DB 1510 ATTCTGAAGGGGAGCTTGGAGTGTGGAAGAGCTCATTTTCTACTACTGCTCATGGCG 1569  
 QY 503 ArgGlyGlnLeuValAlaValGlyLysGlnAsnSerThrMet----- 516  
 DB 1570 CAGGAGACATCATCCAGACTGGAAACCATCTACTCACAGGTGAGCCAGAGAAAGCTCCA 1629  
 QY 517 -----PheSerLeuThrProGluAsnSerTyrProLysAlaCys 530  
 DB 1630 GTAAAGGAATTTGCTTGAGATCCCTGAGAGTTTACGATGATGATGCTCCCATGCTAA 1689  
 QY 531 ValIleValTyrTyrIleGluAspAspGlyGluIleIleSerAspValLeuLysIlePro 550  
 DB 1690 AGTCATCTTACACATCTTGCTGCTGATGGAAGATGATGAGATTCGTGTAACCTTGAA 1749  
 QY 551 ValGlnLeuValPheLysAsnLysIleLysLeuTyrTyrSerLysValLysIleGluPro 570  
 DB 1750 ATTGAAGATGCTCTTGCACAAAGTGACCTGCGCTTACGACATTTCTCAAAGTCTTCC 1809  
 QY 571 SerGluLysValSerLeuArgIleSerValThrGlnProAspSerIleValGlyIleVal 590  
 DB 1810 GCCTCAAAACCCGCTGACAGCTTCAAGCTTCT---CTCACTCCCTCTGTGGCTGAGA 1866  
 QY 591 AlaValAspLysSerValAsnLeuMetCbsnAlaSerAsnAspIleThrMetGluAsnVal 610  
 DB 1867 GCTGTGACACAGAGTGTCTACTCTGTAACCCGAGTGTGAGCTCTCCCTTCC----- 1920  
 QY 611 ValHisGluLeuGlnLeuTyrAsnThrArgLysTyrTyrLeuGlyMetPheMetSerPhe 630  
 DB 1921 -----TGAATATACATCTGCCA-----GGTATGACGACAAACAAATTC 1959  
 QY 631 AlaValPhe-----GlnGluCysGlyLeu-----TyrVal 640  
 DB 1960 GTTCCAGTTCCTCGTCTGTGAAGACCAAGAAAGCTGTATCTGTACAGTTCACTGTTG 2019  
 QY 641 -----Leu 641  
 DB 2020 GCTGAAGACACAAACTTAGTACACATGGAACGTGAAGAGATGTCTATGATATGTG 2079  
 QY 642 ThrAspAlaAsnLeuThrLys----- 648  
 DB 2080 GAGGACATGGGTTTAAACGATTCACCAACTTGATGATCAAACTTCTATAATTTGTTT 2139  
 QY 649 AspTyr----- 650  
 DB 2140 GACTATGGATGGTTCATATACAGACCTCGTGTGAAATTTGATCTTGACCTTACCCCT 2199  
 QY 651 ---IleAsp-GlyValTyrAspAsnAlaGluTyrAlaGluLysPheMetGluGluAsnGly 669  
 DB 2200 GAAATTTCTTGAGATT-----TACGACAAACATGTAGTAAAGACCTGAA 2244  
 QY 669 uclYhisIleValAspIleHisAspPheSerLeuGlySerSerProHisValArgLysHis 689  
 DB 2245 GAGCCACACAGTAAAGATTCATGATGAAATGATCCA-CTTACAGAGACCATCTGAAATA 2303  
 QY 689 sPheProGluThrTyrIleTyrLeuAspThrAsnMetGlyTyrArgIleTyrGlnLysPhe 709

DB 2304 TTTTCTCTAAACCTGGGCTGGATATAGTACAGTAAACTCCACAGACTGCTGGAAGT 2363  
 QY 709 eGluValThrValProAspSerIleThrSerTyrValAlaThrGlyPheValIleSerGly 729  
 DB 2364 GGAATGATCTGTCTTACACATCATCTGAATGGAAGCAGAGAGCTCTCTGTTCCAA 2423  
 QY 729 uAspLeuGlyLeuGlyLeuThrThrThrProValGluLeuGlnAlaPheGlnProPhePhe 749  
 DB 2424 TGACATGCTGCTTGCCCTCTTCTGTG---GTACCTCTCCAAAGCTTCAAGCCCTTCTT 2480  
 QY 749 eIlePheLeuAsnLeuProTyrSerValIleArgGlyGluGluPheAlaLeuGluIleThr 769  
 DB 2481 TGTGAGGCTTCATGTGCTTATCTGTGTGCTGCGTGAAGAACCTTCATGCTCAAGGCCAC 2540  
 QY 769 rIlePheAsnTyrLeuLysAspAlaThrGlnValLysValIleIleGlyLysSerAspLys 789  
 DB 2541 TGTGATGAATCATCTCCACACAGCATGATGATGTGACAGCTGGAAGCCCTCTCTGA 2600  
 QY 789 sPheAspIleLeuMetThrSerSerGluIleAsnAlaThrGlyHisGlnGlnThrLeuLe 809  
 DB 2601 TTTCACACTGTCCACAGGGAAGAC-----CAAGATCTTACTG 2642  
 QY 809 uValProSerGluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGlyGly 829  
 DB 2643 CCGTACGTCACATGGAGGACACACCTCATCTGCTGCTGTAACCTCAAGTCTTAAAGGAA 2702  
 QY 829 uIleProIleThrValThrAla-----LeuSerProThrAlaSerAspAl 844  
 DB 2703 TGTGAATTTCTGTGTGTGCGGAAGACAAACAGTCTCAGAGCCCTGTGTGTTGAGGT 2762  
 QY 844 aValThr-----Gln-MetIleLeuValLysAlaG 854  
 DB 2763 GCCACAGTCTCTGAACCTGGAGAAAGACAGTAAAGTCTGATTAAGTGAAGCTG 2822  
 QY 854 IugIlyIleGluLysSerTyrSerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeu 874  
 DB 2823 AAGGATCAACAGACATGATCATCTGACGTCTGTGATCATGATGCTGAGATAT 2882  
 QY 874 IAsnThrLeuLysThrLeuSerPheSerPheProProAsnThrValThrLysGlyAla 894  
 DB 2883 CT-----GAAAAATGTCCTGCTCTCCACCAACAGTGTGAAGATTCAGCA 2933  
 QY 894 rGValGlnIleThrAlaIleGlyAspValLeuGlyProSerIleAsnGlyLeuLysSerI 914  
 DB 2934 GAGCCATTTCTCTGTGATGAGTGTATCTTGAAGTTCAGCCATTAAGAAACACAAAC 2993  
 QY 914 euIleArgMetProTyrGlyCysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleT 934  
 DB 2994 TTTCATATGCTGCTATGCTGTGGGAGGAGCAACATGTCTTTTGTCTCCAAACATCT 3053  
 QY 934 rYrIleLeuAspTyrLeuThrLysLysGlyGlnLeuThrAspAsnLeuLysGlyLysAla 954  
 DB 3054 ACGTACTGAATATCTGATGATGAACCCAAACGCTGACTCAGAAAGTCAAGCTAAAGCC 3113  
 QY 954 euSerPheMetArgGlyGlyTyrGlnArgGlyLeuLeuTyrGlnArgGluAspGlySerP 974  
 DB 3114 TTGGTCTCTGAGAGCTGTATATGAGAGAGCTGAACTACAAACACAGAGATGCTCTCT 3173  
 QY 974 heSerAlaPheGly-----AsnTyrAspProSerGlySerThrTyrLeuSerAlaPheV 992  
 DB 3174 ACAAGTCTTGGGGATCAAAATGGCCAAAGAAAGAAACACTTGGCTCACAGCTTTG 3233  
 QY 992 aLeuArgCysPheLeuGluAlaAspProTyrIleAspIleAspGlnAsnValLeuHisAla 1012  
 DB 3234 TGCTCAAGTCTTGGCCCAAGCTGAGGCTTCATCTTATGATGAATCAACACATCAACC 3293  
 QY 1012 rGlnTyrThrTyrLeuLysGlyHisGlnLysSerAspGlyGluPheThrAspProGlyAla 1032  
 DB 3294 ATGCTTACCTGCTGCTCCAAAGAGAAAGACAAATGCTGCTTCCGAGCTCCGAT 3353  
 QY 1032 rGValIleHisSerGluLeuGlnGlyLeuLysSerProValThrLeuThrAlaTyrI 1052

```

Db 3354 CATTGTTCAACAATGCGAGGAGGAGTAGATGATGAATGAAATGACCTCTCTGCTTACA 3413
Qy 1052 levalThrSerLeuLeuGlyTyrArgLysPheProAsnIleAspValGlnGlnSer 1072
Db 3414 TAACCATGCGCTTCTGAAAAGTTTCACTCCAGCCAGCATCTGTTGCTCCAAAGCCC 3473
Qy 1072 leHisPheLeuGlnSer-----GlnPheSerArgGlyIleSerAspAsnT 1087
Db 3474 TGACCTGCTGAGCATCTCGAAGACATATAGAAACAAGAACGAATGCGCTTGTGTGT 3533
Qy 1087 TTTTThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGlySerPro---LysAlaLysG 1106
Db 3534 ACACCAAGAGCTTGAATGCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3593
Qy 1106 LysAlaLeuAsnMetLeuThrTyrArgLysGlnGlnGlyGlyMetGlnPheThrVal- 1125
Db 3594 AAATCTGTAATCTCATCTGATGAGAAAGTATATATAAAGAAACAATCTATCCATCGAATA 3653
Qy 1126 --SerSerGlnSerLysLeuSerAspSerTrp-----GlnProArgSerLeuA 1141
Db 3654 GACCTCAAGAAATCCAGAAATCTGAAACACATTTGTACAAACCCAGGCTTCTGCTGTG 3713
Qy 1141 epIleGlnValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGlnG 1161
Db 3714 AAGTAGAGATGAATGATATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1171
Qy 1161 LysIlePro-----IleMetArgTrpLeuSerArg 1171
Db 3762 CAGCCCATCCCTGAGAGATCTGACTTTGTGTAAGAGACATCATGTGCTCAAAAGC 3821
Qy 1171 LysArgAsnSerLeuGlyGlyPheAlaSerThrGlnAspThrThrValAlaLeuLysAla 1191
Db 3822 AACGAAATTCGAATGTGCTTCTCTCCACACAGACATGTGTGTGTGTGTGTGTGTGTGTG 3881
Qy 1191 eusSerGlnPheAlaAlaLeu-----MetAsnThrGlnArgThrAsnIleGlnValT 1208
Db 3882 TGTCCTCAATATGAGACAGTACTTTTTCGAGAAAGTCAAGAAATCTTGTGTGTGTGTGT 3941
Qy 1208 hrValThrGlyProSerSerPro-----SerProLeuAlaValValG 1222
Db 3942 AATCTACAGGGTCACTTTTCCAAAAGTTCCAAAGTGAGAAACAGTAATCGCTGTACTAC 4001
Qy 1222 LysProMetAlaVal-----AsnIleSerAlaAsnGlyPheGlyP 1235
Db 4002 AGCAGGTGCGATTCACAGACATTCCTGGGAGCATATACATCATGTGTGTGTGTGTGTGT 4061
Qy 1235 heaAlaIleCysGlnLeuAsnValValTyrAsnValLysAlaSerGlySerSerArgArg 1255
Db 4062 GTGTGTATGCTCAAGCTATGCTGAGATACAAACATGCACTTGAG----- 4105
Qy 1255 rGArgSerIleGlnAsnGlnGlnLysPheAspLeuAspValAlaValLys----- 1271
Db 4106 -----AAGCAGCTGTCTGCAATTTGCTATATGAGGATCAGACAGTACTTAACTT 4154
Qy 1272 --GluAsnLysAspAspLeuAsnHisValAspLeuAsnValCysThrSerPheSerGlyP 1291
Db 4155 GTAATTAACCCCAAGGCGCAACAAGCTTCCAGATCTCATGAAATGATGATGACAGGCA 4214
Qy 1291 ro---GlyArgSerGlyMetAlaLeuMetGlnValAsnLeuLeuSerGlyPheMetValP 1310
Db 4215 GCCCGCCAGCTTCAACAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4267
Qy 1310 roserGlnAlaIleSerLeuSerGlnThrValLysLysVal----- 1323
Db 4268 -----ATCCCATTTGAACCAACAGTGAAGAAAGCTTGAAGATTTAGAGACAGCTGA 4316
Qy 1324 -----GluTyrAspHisGlyLysLeuAsnLeuTyrLeuAspSerValAsnGlnThrG 1341
Db 4317 GCAGAACAGAAAGTAGAACAACAATGCTTAATATATTTGATGACAGTGAACAATGAGA 4376
Qy 1341 LysPheCysValAsnIleProAlaValArgAsnPheLysValSerAsnThrGlnAspAla 1361
Db 4377 CACTGGCTTCTCTTCAATCATTCACAAAGATATCCAGTAAAGAACCTGAGAGCTGCA 4436

```

```

Qy 1361 erValSerIleValAspTyrTyrGlnProArgGlnAlaValArgSerTyrAsnSer 1380
Db 4437 TTGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4495

RESULT 7
BUS17328
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

```

```

1 (bases 1 to 879)
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
National Institutes of Health, Mammalian Gene Collection (MGC)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.jnl.gov
Plate: LLM414090 Row: P Column: 10
High quality sequence stop: 663.
Location/Qualifiers
1..879
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6514905"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_71"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NciI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.1 kb."

```

```

BASE COUNT      248 a      171 c      206 g      253 t      1 others
ORIGIN
Alignment Scores:
Pred. No.:      1,64e-126      Length:      879
Score:          1280.00      Matches:      263
Percent Similarity: 91.38%      Conservative: 2
Best Local Similarity: 90.69%      Mismatches: 19
Query Match:      17,42%      Gaps:      2
DB:              13

```

```

US-10-020-095-4 (1-1428) x BUS17328 (1-879)
Qy 1070 GlnSerIleHisPheLeuGlnSerGlnPheSerArgGlyIleSerAspAsnTyrThrLeu 1089
Db 2 GACTCTATCAATTTTGTGAGTCTGAATTCAGTAGAGAAATTCACAAATTAATCTTA 61
Qy 1090 AlaLeuIleThrTyrAlaLeuSerSerValGlySerProLysAlaLysGlnAlaLeuAsn 1109
Db 62 GCCCTTAATCAATTTTGTGAGTCTGAATTCAGTAGAGAAATTCACAAATTAATCTTA 121
Qy 1110 MetLeuThrTrpArgLysGlnGlnGlnGlyGlyMetGlnPheThrValSerSerGlnSer 1129
Db 122 ATGCTGATTTGAGAGACAAACAAGAGTGGCATCAATTTCTGAGTGTATGAGAGTCC 181
Qy 1130 LysLeuSerAspSerTrpGlnProArgSerLeuAspIleGlnValAlaAlaTyrAlaLeu 1149
Db 182 AAATCTTGAATCTCTGAGAGACCAAGCTCTCGAGATTAATGAAGTTGACCATATGACATG 241

```





Oy		220	GlnAlaSerGluYrValLeuProLysPheGluValThrIleuGlnThrProLysCys	239
Dd		185	CAGGTTTCAGAAATGTGTTTACCAAAATTGAAGTACTTGCGACACCATTATATCT	126
Oy		240	SerMetAsnSerLysHisIleuAsnGlyThrIleThralalysTythrTygLYlsPro	259
Dd		125	TCTATGAAATCTTAAGCATTTTAAATGGTACCATCACGGCAAAGTATATATGGAGAACA	66
Oy		260	ValIleGlyAapValThrLeuThrPheLeuProLeuserrPhetPGlyLylysLys	278
Dd		65	GTGAAGAGACGTAACCGTTACATTTCCTTTTCCTTTTGGGAAAAAAA	9
RESULT 9				
LOCUS	AL576908/c			
DEFINITION	AL576908 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA			
ACCESSION	clone CSOD1081YB18 3-PRIME, mRNA sequence.			
VERSION	AL576908			
KEYWORDS	AL576908.2 GI:31315187			
SOURCE	EST.			
ORGANISM	Homo sapiens (human)			
REFERENCE	Homo sapiens			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi;			
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
JOURNAL	I. W.B., Gruber,C., Jesssee,J. and Polajcs,D.			
COMMENT	Full-length cDNA libraries and normalization Unpublished On Feb 16, 2001 this sequence version replaced gi:12939516. Contact : Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seque@genoscope.cns.fr Web : www.genoscope.cns.fr Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 3578.f For more information about this cluster, see <a href="http://www.genoscope.cns.fr/">http://www.genoscope.cns.fr/</a> cgi-bin/cluster.cgi?seq=CSOD1081DA09NP&cluster=3578.f. Contact : Feng Jiang Email : fjiang@lifetech.com URL : <a href="http://fulllength.invitrogen.com/">http://fulllength.invitrogen.com/</a> Invitrogen Corporation 1600 Parade Avenue Genoscope sequence ID : CSOD1081DA09NPL. Location/Qualifiers 1..1201 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CSOD1081YB18" /tissue_type="PLACENTA COT 25-NORMALIZED" /clone_id="Homo sapiens PLACENTA COT 25-NORMALIZED" /note="1st strand cDNA was primed with a NotI-clig9(drr) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."			
BASE COUNT	344 a 248 c 224 g 321 t			64 others
ORIGIN				
Alignment Scores:				
Pred. No.:	1 662-98		Length:	1201
Score:	1022.00		Matches:	222
Percent Similarity:	88.84%		Conservative:	1
Best local Similarity:	88.45%		Mismatches:	27
Query Match:	13.91%		Indels:	5
DB:	9		Gaps:	1
US-10-020-095-4 (1-1428) x AL576908 (1-1201)				
Oy	1178 PheaIaserrhGInaspThrhThrValAlaleuLysalaLeuSergLuPheAlaIaleu	1197		
Dd	1030 TTTGCATCTACGCA-GATAACCACT--KMGCTTAARGYCTTYYRA-ATTGCACCCCTA	976		
Oy	1198 MetAnthrGluArgThrAsnIleGlnValThrValThrGlyProSerSerProSerPro	1217		

Db	975	ATGAAMACGAAGG-ACAAATATCCAGTAGCAGCGGGCTT-AGCTCACCAGTCT	918
Qy	1218	LeuAlaValValGlnProMetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIle	1237
Db	917	CTGGCTGTGGAAACAGCCACGCGAGTTAAATATTCGCAATGGATTTTGATTTGCTATT	858
Qy	1238	CyGGIleuAsnValValTyrAsnValLysAlaSerGlySerSerArgArgArgSer	1257
Db	857	TGTACGCTCAATGTTGTATATTAATGTGAGAGGCTTTCGGCTCTCTAGAGACGAAATCT	798
Qy	1258	IIeGIIAsnGlnGluAlaPheAspLeuAspValAlaValIysGlnuAsnLysAspAspLeu	1277
Db	797	ATCCAAATATCAGAGGCTTTGATTATAGATGTCGCTGTAAAGAAATTAAGATGATCTC	738
Qy	1278	AsnHisValAspLeuAsnValCysIleHisPheSerGlyProGlyArgSerGlyMetAla	1297
Db	737	AATCATGTGATTTGGATGTGTGACAAACGTTTTTCGGGCCCGGGTGAAGTGGCATGGCT	678
Qy	1298	LeuMetGlnValAsnLeuLeuSerGlyPheMetValProSerGluAlaIleSerLeuSer	1317
Db	677	CTTATGGAAGTTAACTTATTAAAGTGGCTTTATGTGCTTACAGAGCAATTTCTTGAGC	618
Qy	1318	GluIleValLysLysValGluTyrAspHisGlyLysLeuAsnLeuTyrLeuAspSerVal	1337
Db	617	GAGACGCTGAAGAAAGGATATGATCATGAAAACTCAACCTCTATTATGATTCGTA	558
Qy	1338	AsnGluIleThrGlnPheCysValAsnIleProAlaValArgAsnPheLysValSerAsnThr	1357
Db	557	AATGAAACCCAGTTTGTGTAAATATCTCGCTGTGAAACCTTTAAAGTTCAAATAC	498
Qy	1358	GluAspAlaSerValSerIleValAspTyrTyrGluProArgArgGluAlaValArgSer	1377
Db	497	CAAGAGGCTTCAGTGTCCATAGTGAGATACATGAGCAAGAGACAGGGCGTGAGAACT	438
Qy	1378	TyrAsnSerGluValLysLeuSerSerCysAspLeuCysSerAspValGlnGlyCysArg	1397
Db	437	AACAACTCTGAAGTGAACACTGCCCCCTGTGACCTTGCGAGTGATGMCAGVGMSCCGM	378
Qy	1398	ProCysGluAspGlyAlaSerGlySerHisHisIleSerSerValIlePheIlePheCys	1417
Db	377	CCTTGATGAGATGAGGCTCAVGGMCCCATCATCATCCTCAGTCATGATTTATTTTCGT	318
Qy	1418	PheLysLeuLeuTyrPheMetGluLeuTyrLeu	1428
Db	317	CTCAGAGCTCTGAACCTTATGGAACCTTGGCTG	285
RESULT 10			
AK042169			
LOCUS			
DEFINITION	AK042169	1495 bp	mRNA linear HTC 05-DEC-2002
	Mus musculus 3 days neonate thymus cDNA, RIKEN full-length enriched		
	library, clone:1630065N19 product:weakly similar to		
	alpha-2-macroglobulin (fragments) [Limulus polyphemus], full insert		
	sequence.		
ACCESSION	AK042169		
VERSION	AK042169.1	GI:26334978	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE			
AUTHORS	Carninci, P. and Hayashizaki, Y.		
TITLE	High-efficiency full-length cDNA cloning		
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)		
MEDLINE	99279253		
PUBMED	10349636		
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,		
	Itoh, M., Komo, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to		
	prepare full-length cDNA libraries for rapid discovery of new genes		
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)		
MEDLINE	20493734		

PUBMED 11042159  
REFERENCE  
AUTHORS  
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,  
Kono,H., Akiyama,J., Nishi,K., Katsunai,T., Tashiro,H., Itoh,M.,  
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,  
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,  
Fujitake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,B., Watanabe,M.,  
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J.,  
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.  
RIKEN integrated sequence analysis (RISA) system-384-format  
sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
4  
Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,  
Aizawa,T., Hara,A., Fukunishi,Y., Kono,H., Adachi,J., Fukuda,S.,  
Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamada,T.,  
Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,  
Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,  
Fleischmann,W., Gaasterland,T., Glass,C., King,B., Kochiwa,H.,  
Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G.,  
Quackenbush,J., Schriml,L.M., Staahl,F., Suzuki,R., Tomita,M.,  
Wagner,L., Washio,T., Sakai,K., Okido,T., Furumori,M., Aono,H.,  
Balderelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,  
Carninci,P., de Bonaudo,M.F., Brownstein,M.D., Bult,C.,  
Fleischer,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,  
Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,  
Marchionni,L., Mashima,T., Mazzarelli,J., Mombaerts,P., Nordone,P.,  
Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,  
Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,  
Toyo-oka,K., Wang,K.H., Weitz,C., Whitaker,C., Wilming,L.,  
Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawai,J., Kohetsuki,S.  
and Hayashizaki,Y.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409 (6821), 685-690 (2001)

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
5  
The FANTOM Consortium and the RIKEN Genome Exploration Research  
Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation  
of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
6  
(bases 1 to 1495)  
Adachi,J., Aizawa,K., Akimura,T., Arikawa,T., Bono,H., Carninci,P.,  
Fukuda,S., Furumori,M., Hanagaki,T., Hara,A., Hashizume,M.,  
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,  
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,  
Katch,H., Kawai,J., Kojima,Y., Kondo,S., Kono,H., Konda,M.,  
Koyama,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,  
Nakamura,M., Nishi,K., Nomura,K., Nunazaki,R., Ohno,M., Ohnato,N.,  
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,  
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,  
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akanita,S.,  
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,  
Muramatsu,M. and Hayashizaki,Y.  
Direct Submission  
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
Kanagawa 230-0045, Japan (E-mail:genome-res@sc.riken.go.jp,  
URL: <http://genome.gsc.riken.go.jp/>, Tel:81-45-503-9222,  
Fax:81-45-503-9216)  
cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.  
Tissues were provided by Dr. John Todd (Dept. of Medical Genetics  
Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome  
Trust/MRC building Addenbrookes Hospital Cambridge) whose

assistance we gratefully acknowledge.  
Please visit our web site for further details.  
URL: <http://genome.gsc.riken.go.jp/>  
URL: <http://fantom.gsc.riken.go.jp/>

## FEATURES

## source

1. 1495

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C57BL/6J"

/db\_xref="FANTOM DB:A630065N19"

/db\_xref="taxon:10090"

/clone="A630065N19"

/tissue\_type="thymus"

/clone\_lib="RIKEN full-length enriched mouse cDNA library"

/dev\_stage="3 days neonate"

121..954

/note="unamed protein product; putative

weakly similar to alpha-2-macroglobulin (fragments)

[limulus polyphemus] (PIR:A5260, evidence: FASTY,

51.8%id, 76.7%length, match=168)"

/codon\_start=1

/protein\_id="BAC31190.1"

/db\_xref="GI:26334979"

/translation="MRWLIQPNLSIGFVSTQDTYVALKALSEFALVHKENTDIQLT

VTGPDIPIPSIHRIDSDMLFIHQEELHALDPIYVVAHSGFALCOLANDYVNGS

ASSDKRRSTENQVEDLDIVNNEDIDSHLNVCTSHLSGSRGTMTLMEVLISGFS

AASDIPLESTLKVKEVNGKLVLDVSNESQFVNIPTVRDYKVNIRDSVAVMD

YYERPRQAVRNVYVYKLSLSCLSPDNCKHTBDATSLARSSSLVFCVLYLTVQ

H"

BASE COUNT 368 a 349 c 334 g 444 t  
ORIGIN

## Alignment Scores:

Pred. No.: 9 39e-97 Length: 1495  
Score: 1007.00 Matches: 202  
Percent Similarity: 76.18% Conservative: 41  
Best Local Similarity: 63.32% Mismatches: 56  
Query Match: 13.70% Indels: 20  
DB: 11 Gaps: 4

US-10-020-095-4 (1-1428) x AK042169 (1-1495)

QY 1124 TTPVALSERGSLUSERLYEUSERASPSERTTPGILNPROARSERLEUAPILLEGIL 1143  
DB 1 TGGTGTGATCAGACGACCGCGCTCTCTGGGTCCTGGCAGCGCGCTCCGTGACATCGA 60  
QY 1144 VALAIAIYRYALEULEUSERHSIHPHEULEGINPNEGINTHSERGLUGIYLEPO 1163  
DB 61 ATCGCTGCTTATGACACTGCGCAGCACTCTGCAC---CATGTGCTGAGGGAATCCCG 117  
QY 1164 ILEMELATGTPLEUSERARGINARGENSERLGUGIYGYPHALASERTHRLINER 1183  
DB 118 GTTAAGAGTGCTCATCCAGCAAGAAACAGCCGTGGAGGTTTGTATCCACTAGAT 177  
QY 1184 THRTRVALAIALEULEUSERGILUPHEAIAIALEULEMERASTHRLGILNRYTH 1203  
DB 178 ACTGTGTGGCCCTTAAGACCATATCTGAATTTTACGCCCTTGTGCACAGAAATAACA 237  
QY 1204 AENILEGINVALTHRVATHRGLYPROSERSEPRO----- 1215  
DB 238 GATATTCACAACTGACCGTACAGAGGCCCGCATCCCATCATCACTTCGGAATTGAC 297  
QY 1216 -----serProLeuAIAValGILNPrOmELaVal 1226  
DB 298 TCCCAAACTCTTCTCCTTCACAGAGAAGCTTACGCGCTGATCCATCAACGTT 357  
QY 1227 AENILESERALAENGLYPHEGLYPHEAIAIAIEGYNLEUASNVAIAIYRASNVAI 1246  
DB 358 AATGTTCTGCGCAGCTCGGAGATTGCTATTTCACAGCTTAATGTTGACATTAACGG 417  
QY 1247 LYSALASERGLYSESRARGTRGTRGTRGTRGTRGTRGTRGTRGTRGTRGTRGTRG 1266  
DB 418 AAAAGTTGAGTTCTTTAAAGACGAGATCTACCAAAATCAAGAGTTTGTATTTA 477

Oy		1267	AspValAlaValLysGluAsnIleLysAspSerLeuAsnHisValAlaSerLeuAsnValCysThr	1286
Db		478	GAGCTCATTTGG---AATTAATGAGACACACTTAACTCACTCGGAATCTTAATGTGTGCACA	534
Oy		1287	SerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluValAsnLeuLeuSerGly	1306
Db		535	AGTCACCTGGGTTCCAGAGAAGACAGGACGTGGCTTATTGGAAGTGAACTTCTTCAGTGGC	594
Oy		1307	PheMetValProSerGluAlaIleSerLeuSerGluTrpValLysValGluTrpAsp	1326
Db		595	TTTAGTGCACTTCAGATTCATTCTCTTGAGTGAACCCTGAAGAAAAGTGAATATGAT	654
Oy		1327	HisGlyLysLeuAsnLeuTyrlenAspSerValAsnGluTrpGlnPheCysValAsnIle	1346
Db		655	AACGGAAACTCACACCTTATTATTGATCTCTGAATGAATCCCACTTTGTGTTAAATTT	714
Oy		1347	ProAlaValArgAsnPhelLysValSerAsnThrGlnAspAlaSerValSerIleValAsp	1366
Db		715	CCCCACTGGAGAGACTACAAAGTTTCCAATTTCCAGATGGTTCAGTATCTGTATATGAC	774
Oy		1367	TyrTyrgluproArgArgGlnAlaValArgSerTyraAsnserGluValLysLeuSerSer	1386
Db		775	TACTATGACCAAGAGGACGAGCGAGTGGCAAGCTAACACACCAGGTAAACTGTCTCTCA	834
Oy		1387	CysAspLeuCySserAspValGInGlyCysArtpProCysGluAspGlyAlaSerGlySer	1406
Db		835	TGTTATCACTCAATCCAGACACC---AAGTCGAAGTACACACAGGACGAGCCACGCACTCC	891
Oy		1407	HisHisSerSerValIlePheIlePheCysPheLysLeuLeuTyrlPheMetGlu	1425
Db		892	CTTGACCGTCTTCGAGCGCTTCGTGCTTTGCTGCCGCTTCGTGTAATTTGTGCAA	948
RESULT 11				
LOCUS	BX386457	978 bp	mRNA	EST 08-MAY-2003
DEFINITION	BX386457 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA			
ACCESSION	clone CS001081YB18 5-PRIME, mRNA sequence.			
VERSION	BX386457			
KEYWORDS	BX386457.1 GI:30457430			
SOURCE	EST.			
ORGANISM	Homo sapiens (human)			
	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	1 (bases 1 to 978)			
AUTHORS	Ii W.B., Gruber C., Jesse J. and Polayes D.			
JOURNAL	Full-length cDNA libraries and normalization			
COMMENT	Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 Evry cedex - France Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr Library was constructed by Life Technologists, a division of Invitrogen. This sequence belongs to sequence cluster 3578.f For more information about this cluster, see http://www.genoscope.cns.fr/ cgl-bin/cluster.cgi?seq=CS1A1021B100Plac1cluster=3578.f. Contact : Feng Liang Email : fliang@life.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Paradey Avenue Genoscope sequence ID : CS1A1021B100Pl.			

BASE COUNT	374 a	151 c	179 g	220 t	54 others
ORIGIN					
Alignment Scores:					
Pred. No.:	6.09e-97	Length:		978	
Score:	1006.00	Matches:		217	
Percent Similarity:	75.59%	Conservative:		9	
Best Local Similarity:	72.58%	Mismatches:		72	
Query Match:	13.69%	Indels:		2	
DB:	13	Gaps:		0	
US-10-020-095-4 (1-1428) x BX36457 (1-978)					
QY	3	GlyProProleuLeuThrAlaAlaHisIleuLeuCyValCysThrAlaAlaLeuAlaVal	22		
DB	84	GGGCGANCGCTCTGTGAACGNAGCAAAATCTCTGGGTGCACCGCGCGTGGCGGTG	143		
QY	23	AlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGlyGlyAsnVal	42		
DB	144	GCTCCAGGCGMTCGGTTTMTGGTGAMAGCCAGGAAATATMAGCAGCAGGAAATGTG	203		
QY	43	ThrIleGlyValGluLeuLeuGluHisCysProSerGlnValThrValValaGluLeu	62		
DB	204	ACWATGGGGGTGAGCTTCTCGAACACTGCMCTTCACAGTGAGTGAAGGCGGA-SWG	262		
QY	63	LeuIleThrAlaSerAsnLeuThrValSerValLeuGluValaGluGlyValaPheGluLys	82		
DB	263	CTCAAGCAGCAGCAAAAAATMAAGTATCAGTAATGAGAAAGAAAGATTTTGA	322		
QY	83	GlySerPheIleThrLeuThrLeuProLeuAsnSerAlaAspGluIleTyr	102		
DB	323	AGATCTTTTAAAAAAAATAATCAAAAAAAMAAAAACAGTCAGATGATATAT	382		
QY	103	GluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThrArgLeu	122		
DB	383	AARCAAGCTAAAGAGAGTACCCAGGATGARATWTATCTCTAAMAGAACAGCMTA	442		
QY	123	SerPheGluThrIleArgIleSerValPheIleGlnThrAspValaLeuTyrLysPro	142		
DB	443	AAATTTGAAACCAAAAAATMAAGAAWTAAAAAAMAAAAAGAAATTAACACCA	502		
QY	143	LysGlnGluValaLysPheArgIleValThrLeuPheSerAspPheLysProTyrLysThr	162		
DB	503	AAGCAAGAGTGAAGTTTGGCTTGTATGACTCTTCTAARAATTAAGCATTAACAAAA	562		
QY	163	SerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnIntProLysGln	182		
DB	563	ACTTTAAACATTATCTTAAGAAACCAATTAATTTGATCCAAACGTGGTGTCAAA	622		
QY	183	GlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisPheIleLeuGly	202		
DB	623	CAAGAGATCTTGGATCTATTCCAAACTTTGACGTATCTTCCATCCAAATCTTGGT	682		
QY	203	AspTrpSerIleGlnValaGlnAlaAsnAspGlnThrTyrGlnAsnPheGlnValSer	222		
DB	683	GACGGCTATTAAGTCTCAAGTGAATGACGACATATATATCAATCATTTCCAGTTCA	742		
QY	223	GluTyrValLeuProLysPheGluValaThrLeuGlnThrProLeuTyrCysSerMetAsn	242		
DB	743	GAATATGATTAACAAATTTGAATGACTTTGCGACACACATTAATTTGTGTATGAAT	802		
QY	243	SerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProValValGly	262		
DB	803	TCTAAGCATTTAAATGTACTCATTTCCAAARSATATAGATTSTAAGTYAAGYAAATCA	862		
QY	263	AspValThrLeuThrPheLeuProLeuSerPheTrpGlyLys-LysLysAsnIleThrLys	282		
DB	863	GACTTAACTCTACATTTTACCTTTTTCCKTGTGGGSSAAKAKAKAAAAATTTTACAA	922		
QY	282	SerThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluGluMetLys	300		
DB	923	AAACATTTAAGATTAATGATCTGCAAAATTTCTCTTTTATATGATGAAGATGAAA	977		

```

RESULT 12
Bg173799          836 bp      mRNA      linear      EST 06-FEB-2001
LOCUS             60233853f1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:4457172 5',
DEFINITION        mRNA sequence.
ACCESSION         Bg173799
VERSION           Bg173799.1 GI:12680502
KEYWORDS          EST.
SOURCE            Mus musculus (house mouse)
ORGANISM          Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE         1 (bases 1 to 836)
AUTHORS           NIH-MGC http://mgs.nci.nih.gov/.
TITLE             National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL           Unpublished
COMMENT           Contact: Robert Strausberg, Ph.D.
                  Email: cgapbs-remail.nih.gov
                  Tissue Procurement: Gilbert Smith, Ph.D.
                  cDNA Library Preparation: Life Technologies, Inc.
                  DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
                  Clone distribution: MGC clone distribution information can be
                  found through the I.M.A.G.E. Consortium/LNL at:
                  http://image.llnl.gov
                  Plate: L1AM10253 row: 1 column: 13
                  High quality sequence stop: 694.
                  Location/Qualifiers
FEATURES
    source
        1..836
            /organism="Mus musculus"
            /mol_type="mRNA"
            /strain="FVB/N"
            /db_xref="taxon:10090"
            /clone="IMAGE:4457172"
            /tissue_type="tumor, biopsy sample"
            /dev_stage="10 months, virgin"
            /lab_host="DH10B"
            /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
            Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
            library constructed by Life Technologies. Investigator
            providing samples: Gilbert Smith, NIH"
BASE COUNT        228 a 179 c 202 g 227 t
ORIGIN
Alignment Scores:
Pred. No.:        3,376-94      Length:      836
Score:            979.50        Matches:     200
Percent Similarity: 87.55%      Conservative: 25
Best Local Similarity: 77.82%   Mismatches:  30
Query Match:      13.33%        Indels:       4
                                Gaps:         1
                                10
US-10-020-095-4 (1-1428) x Bg173799 (1-836)
QY      942  LysLysGlnLeuThrAspAsnLeuYsgLysAlaLeuSerPheMetArgGlnGlyTyr 961
Db      1042 AenLysSerProValThrLeuThrAlaTyrLleValThrSerLeuLeuGlyTyrArgLys 1061
        1  CAGAAACGCTGACAGCTTAATTTAAAGAAAGGCCCTTCAATATATGAGGCAAGCTTAC 60
        301 ACCAAAGCCAGTAACCTTACGACCTTACGACCTTATGATGTTGCTGGGACACAAAG 360
QY      962  GlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyrAsp 981
Db      1062 TyrglnProAenLleAspValGlnGluSerLleHisPheLeuGluSerArg 1081
        61  CAAAGGAGGCTTCTATACAGAGGAGAGTGCTCTTCAAGTCTTTGGGAGCAATTGAC 120
        361 TATCAGCCTATATCATGATGATCAAGACCTCAATCAAGTTTGG- GAATTTGAATTCAGCAGA 419
QY      982  ProSerGlySerThrTTrpLeuSerAlaPheValLeuArgCysPheLeuGluAlaAspPro 1001
Db      1082 GlyLleSerAspAsnTyrThrLeuAlaLeuLleThrTyrAlaLeuSerSerValGlySer 1101
        121 TCTTGTGGAGACACTTGCTGTACACATTTGTTTAAAGTGTCTTCTCGAAGCTGATTAC 180
        420 GGAATTTGCGACAATTATACCTTAGCAATTATATCCATAGCCCTGTCCACAGTCGGAGC 479
QY      1002 TyrlleAspLleAspGlnAsnValLeuHisArgThrTyrThrTrpLeuLysGlyHisGln 1021
Db      1102 TyrlleAspLleAspGlnAsnValLeuHisArgThrTyrThrTrpLeuLysGlyHisGln 1041
        181 TATATAGATATTGATCAGGATGTGTACACAAACATATCTGCTTAATGACATAG 240
        479 1022 LysSerAsnGlyGluPheTrpAspProGlyArgValLleHisSerGluLeuGlnGly 1041
        1041

```

```

Db      241  AATTCATAGTCGATTTTGGAGCCAGAAAGACTGATTACAGTGAACTTCAAGTGGC 300
QY      1042 AenLysSerProValThrLeuThrAlaTyrLleValThrSerLeuLeuGlyTyrArgLys 1061
Db      301 ACCAAAGCCAGTAACCTTACGACCTTACGACCTTATGATGTTGCTGGGACACAAAG 360
QY      1062 TyrglnProAenLleAspValGlnGluSerLleHisPheLeuGluSerArg 1081
Db      361 TATCAGCCTATATCATGATGATCAAGACCTCAATCAAGTTTGG- GAATTTGAATTCAGCAGA 419
QY      1082 GlyLleSerAspAsnTyrThrLeuAlaLeuLleThrTyrAlaLeuSerSerValGlySer 1101
Db      420 GGAATTTGCGACAATTATACCTTAGCAATTATATCCATAGCCCTGTCCACAGTCGGAGC 479
QY      1102 ProLysAlaLysGluAlaLeuAsnMetLeuThrTrpArgAlaGlnGlnGlyGlyMet 1121
Db      480 CTTAAAGAGAGAGGCTTGAACCTGCTGATGACAGCATGAGAAAGAGAGAGACACA 539
QY      1122 GlnPheTrpValSerSerGluSerLysLeuSerAspSerTrpGlnProArgSerLeuAsp 1141
Db      540 CAGTTCTGTGTATCATAGACCTGCGCTCTCTGCGTCTGCGACGCGCTCGGTGAC 599
QY      1142 lleGluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGlnGly 1161
Db      600 ATCGAAATCGCTTATGCACTGTGTGCGCACACTGTGCAC--CATGTCTTGAAGA 656
QY      1162 lleProLleMetArgTrpLeuSerArgGlnArgAsnSerLeuGlyGlyPheAlaSerTh 1181
Db      657 ATCCCGGTATGAAAGTGCTTCATCCAG3CAAAGAAAGCGCTGGAGG-TTTGTAATCCAC 715
QY      1181 rgLAspThrThrValAlaLeuLysAlaLeuSerGluPheAlaAlaLeu 1197
Db      716 TCAGATATTCTTGTGGGCTTGAAGGCAATTTTCTGAATTTTCCGCTTG 764
RESULT 13
B0624356/c       638 bp      mRNA      linear      EST 23-SEP-2002
LOCUS             UI-H-FG1-bgj-1-22-0-UI.s1 NCI CGAP_FG1 Homo sapiens cDNA clone
DEFINITION        UI-H-FG1-bgj-1-22-0-UI 3', mRNA sequence.
ACCESSION         B0624356
VERSION           B0624356.1 GI:23290571
KEYWORDS          EST.
SOURCE            Homo sapiens (human)
ORGANISM          Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE         1 (bases 1 to 638)
AUTHORS           NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE             National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
                  Tumor Gene Index
JOURNAL           Unpublished
COMMENT           Contact: Robert Strausberg, Ph.D.
                  Email: cgapbs-remail.nih.gov
                  Tissue Procurement: James Martin
                  cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
                  DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
                  Clone Distribution: Clone distribution information can be obtained
                  from Dr. M. Bento Soares, bento-soares@uiowa.edu
                  Seq primer: M13 FORWARD
FEATURES
    source
        1..638
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="UI-H-FG1-bgj-1-22-0-UI"
            /tissue_type="Cell lines"
            /dev_stage="Adult"
            /lab_host="DH10B (Life Technologies)"
            /clone_lib="NCI CGAP FG1"
            /note="Organ: Enchondroma; Vector: pT73-Pac (Pharmacia)
            with a modified polylinker; Site_1: Bcor I; Site_2: Not I;

```



prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 100.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATTCCTGATTATTAATATCCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified phlucscript KS(+) after bulk excision from lambda FLC I."

BASE COUNT 195 a 118 c 140 g 194 t 1 others  
ORIGIN

## Alignment Scores:

Pred. No.: 3,12e-90 Length: 648  
Score: 941.00 Matches: 178  
Percent Similarity: 92.09% Conservative: 20  
Best Local Similarity: 82.79% Mismatches: 17  
Query Match: 12.81% Indels: 0  
DB: 10 Gaps: 0

US-10-020-095-4 (1-1428) x BB613975 (1-648)

QY 862 GlnSerlleuLeuAspLeuThrAspAsnArgLeuGlnSerThrleuYsThrleuSer 881  
DB 2 AATCTGCTTATGGATCTGACAGATGCAACGTAAGAAAGTAGACCAATCTATGACA 61  
QY 882 PheSerPheProAsnThrValThrGlySerGluArgValGlnleThrAlaIleGly 901  
DB 62 TTTCTCTTCTCCGATGAGGATGCGATGCGAAGTAAGTTCAGATCCAGCAATGGA 121  
QY 902 AspValleuGlyProSerlleuAsnGlyLeuAlaSerlleuIleArgMetProTyrGlyCys 921  
DB 122 GATATCCTTGTTCCCTCCATCAATGCTTATCTTCACTGATCCGATGCTTATGATGT 181  
QY 922 GlyGluGlnAsnMetlleAsnPhelaProAsnIleTyrIleLeuAspTyrleuThrIys 941  
DB 182 GGTGAACAGAACATGATATATTTGCTCCAAATATTTACATTTGGATTATCTGACTAA 241  
QY 942 LysLysGlnLeuThrAspAsnleuYsGluValAlaLeuSerPheMetArgGlnGlyTyr 961  
DB 242 CAGAAACAGCTGACGATTATTAAGAAAGGCCCTTCAATATGAGGCAAGGTTC 301  
QY 962 GlnArgGluLeuLeuTyrGlnArgGlySerPheSerAlaPheGlyAsnTyrAsp 981  
DB 302 CAAAGGAGCTTCTCATCAGAGGAGAGTGGCTCTTCAAGCTTTGGGACATGAC 361  
QY 982 ProSerGlySerThrTyrleuSerAlaPheValleuArgCysPheleuGlnAlaAspPro 1001  
DB 362 TCTTCTGGGAGACCTTGCTGACGATTGTTTAAAGATGCTTCTGGAACCTGATTAC 421  
QY 1002 TyrIleAspIleAspGlnAsnValleuHISargThrTyrThrTyrleuYsGlnIsgln 1021  
DB 422 TATATAGATATGATGAGATGTTTACACAGAACTATATCTGGCTTATATGACACATAG 481  
QY 1022 LysSerAsnGlyGluPheTyrAspProGlyArgValIleHISerGluLeuGlnGlyI 1041  
DB 482 AATTCATAGCTGATTTTGGAGCCAGAGAGATGATTCACAGTCACTTCAAGGTGC 541  
QY 1042 AsnLysSerProValThrleuThrAlaTyrIleValThrSerleuLeuGlyTyrArgIys 1061  
DB 542 ACCAAAGGCCAGTAACTTACGGGCTATATGTTGCTTCTGCTGGATATACAAAG 601  
QY 1062 TyrGlnProAsnIleAspValGlnGlnSerIleHISerPheleuGln 1076  
DB 602 TATCAGCCTAATATGATGATGATCAACATCAAGTTTGGAA 646

## RESULT 15

AA418644

LOCUS

DEFINITION AA418644 634 bp mRNA linear EST 16-OCT-1997

similar to TR:G534873 G534873 ALPHA-2-MACROGLOBULIN, ;, mRNA

ACCESSION

AA418644

VERSION AA418644.1 GI:2080463  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 634)  
Hillier, L., Allam, M., Bowles, L., Dubuque, T., Geisler, G., Jost, S., Kuwaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Stepien, M., Tan, F., Theisling, B., White, Y., Wyllie, T., Waterston, R., and Wilson, R.  
WashU-Merck EST Project 1997

TITLE JOURNAL  
COMMENT Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu

This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.  
Seq primer: -28m13 rev2 ET from Amersham  
High quality sequence stop: 494.

## FEATURES

source

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:767365"  
/tissue\_type="Pooled human melanocyte, fetal heart, and pregnant uterus"  
/lab\_host="DH10B"  
/clone\_lib="Soares NBHMPu S1"  
/note="Organ: mixed (see below); Vector: pTTT3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not 1; Site 2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (melanocyte 2NBH, pregnant heart NBHPU, and fetal heart NBH19W) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of 1 M.A.G.E. clones 260232-265223, 340488-345479, and 484488-489479."

BASE COUNT 174 a 109 c 139 g 212 t

ORIGIN

## Alignment Scores:

Pred. No.: 3.99e-87 Length: 634  
Score: 912.00 Matches: 174  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 12.41% Indels: 0  
DB: 9 Gaps: 0

US-10-020-095-4 (1-1428) x AA418644 (1-634)

QY 1255 ArgArgSerIleGlnAsnGlnAlaPheAspLeuAspValAlaValIysGlnAsnLys 1274  
DB 1 CGAAGATCTATCCAAATCAAGAAAGCCTTGTATTAGATGTTGGCTGTAAAGAAATATAA 60  
QY 1275 AspAspLeuAsnHISValAspLeuAsnValCysThrSerPheSerGlyProGlyArgSer 1294  
DB 61 GATGATCTCAATCAATGAGATTTGAATGTGTACACAGCTTTTGGGGCCCGGATGAGGT 120  
QY 1295 GlyMetAlaLeuMetGluValAsnLeuLeuSerGlyPheMetValProSerGluAlaIle 1314  
DB 121 GGCATGGCTCTTATGGAAGTAACTTAACTTAAGTGGCTTAAAGTGCCCTTCAAGAACATT 180  
QY 1315 SerLeuSerGluThrValIleValIysValGluTyrAspHISGlyIleLeuAsnLeuTyrleu 1334  
DB 181 TCTTGACGACGACAGTGAAGAAAGTGAATATGATCATGAGAAACATCAACCTCTATTTA 240



```
Qy 1335 AspSerValaengluThrglnPheCyValaenilleProalaValaArgaenPheLyVal 1354
    |||
Db 241 GATTCTGTAAATGAAGCCAGTTTGTGTAAATATCTCTGCTGAGAGAACTTTAAAGTT 300
    |||
Qy 1355 SerAsnThrGlnAspAlaSerValSerilleValaAspTyrTyrGlnProArgArgGlnAla 1374
    |||
Db 301 TCAATATCCCAAGATGCTTCAAGTGTCTCAATGATTAATGAGCCAGAGAGACAGGCG 360
    |||
Qy 1375 ValArgSerTyrAsnSerGluValIlyleuSerSerCyAspLeuCySerAspValGln 1394
    |||
Db 361 GTGAGAGATTACAACTGAGAGTGAAGCTGTCTCTGTGACCTTTGAGATGATGTCAG 420
    |||
Qy 1395 GlyCySarGProCyGluAspGlyAlaSerGlySerHisHisSerSerValIlePhe 1414
    |||
Db 421 GGCTGCCGTCTGTGAGGATGAGAGCTTCCATCATCACTCTTCAGTCATTTT 480
    |||
Qy 1415 IlePheCyPheLyLeuLeuTyrPheMetGluLeuTyrPleu 1428
    |||
Db 481 ATTTCGTGTCAAGCTTGTGACTTATGSACTTGGCTG 522
    |||
```

Search completed: January 17, 2004, 02:52:00  
Job time : 5723 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2004, 18:02:49 ; Search time 54 Seconds  
(without alignments)  
4197.436 Million cell updates/sec

Title: US-10-020-095-4

Perfect score: 7348  
Sequence: 1 MGGPPLTAHLICVCTNAL.....HSSVIFFCFLYFMEWL 1428

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

A.Geneseq\_19Jun03.\*  
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7331	99.8	1428	AAE32012	Human r150 protein
2	7326.5	99.7	1445	ABB82165	Human CD109 KI pro
3	7326.5	99.7	1445	ABB82167	Human CD109 KI-H7
4	7326.5	99.7	1445	AAO19372	Human platelet all
5	7318.5	99.6	1445	AAE32013	Human r150 protein
6	7317.5	99.6	1445	ABB82166	Human CD109 KI var
7	7317.5	99.6	1445	ABB82168	Human CD109 KI-H7
8	7317.5	99.6	1445	AAO19373	Human platelet all
9	7307.5	99.4	1445	AAE12127	Hydrophobic domain

10	7081	96.4	1382	24	AAE32014	Human blood cell s
11	7068.5	96.2	1399	24	AAE32015	Human r150 protein
12	5470	74.4	1067	23	ABB82169	Human CD109 KI5 pr
13	5461	74.3	1067	23	ABB82170	Human CD109 KI5 va
14	2860.5	38.9	1445	24	AAE32018	Human blood cell s
15	1854.5	25.2	1508	22	AAE30676	C. elegans alpha-2
16	1850	25.2	1519	23	AAE50677	C. elegans alpha-2
17	1628.5	22.2	336	23	AAE90069	Human polypeptide
18	1592	21.7	352	22	AAE11581	Human secreted pro
19	1441	19.6	1474	21	AAE97157	Human alpha-2-macr
20	1440	19.6	1474	22	AAE50673	Human alpha-2 macr
21	1440	19.6	1474	23	AAE5218	Hypoxia-regulated
22	1440	19.6	1474	23	AAU74798	Human alpha 2 macr
23	1440	19.6	1474	23	AAU81017	Human alpha2 macro
24	1437.5	19.6	1508	22	AAE63548	A human alpha-2 ma
25	1437	19.6	1451	23	AAU81018	Human alpha2 macro
26	1430	19.5	1203	22	AAE66669	Drosophila melanog
27	1428.5	19.4	1474	12	AAE11334	Recombinant human
28	1425	19.4	1484	12	AAE11749	Human alpha-2 macr
29	1418.5	19.3	1492	23	ABE76898	Human alpha-2-macr
30	1413.5	19.2	1492	23	ABE61768	Novel alpha-2-macr
31	1406	19.1	1425	22	ABE71908	Drosophila melanog
32	1396.5	19.0	1487	22	ABE28006	Novel human diagno
33	1376.5	18.7	1479	23	ABE09519	Human pregnancy zo
34	1352.5	18.4	1475	22	ABE21224	Novel human diagno
35	1329.5	18.1	281	23	ABE93092	Alpha-1 proteinase
36	1328.5	18.1	1760	22	ABE63367	Drosophila melanog
37	1328.5	18.1	1760	23	ABE70020	Larval viability a
38	1315.5	17.9	1342	22	ABE28005	Novel human diagno
39	1299.5	17.7	1413	22	ABE62921	Drosophila melanog
40	1255	17.1	1285	21	AAE43949	Human cancer assoc
41	1173.5	16.0	248	22	AAE74689	Human pro tease and
42	1157	15.7	1411	24	ABU52372	Human GPCR related
43	1109.5	15.1	1436	24	ABU52373	Human GPCR related
44	1101.5	15.0	753	22	ABE67122	Drosophila melanog
45	1099.5	15.0	798	22	ABE62893	Drosophila melanog

#### ALIGNMENTS

RESULT 1	AAE32012	AAE32012 standard; Protein; 1428 AA.
ID	AAE32012	
AC	AAE32012	
XX		
DT	24-MAR-2003	(first entry)
XX		
DE	Human r150 protein #1.	
XX		
KW	Human; GPI-anchored TGF-beta1 binding protein; r150 protein; cancer;	
KW	glycosylphosphatidyl inositol; transforming growth factor-beta1;	
KW	therapy.	
XX		
OS	Homo sapiens.	
XX		
PH		
FT	Key	Location/Qualifiers
FT	Misc-difference	703
FT	/label=	Ser, Tyr
FT	/note=	"Encoded by TNC"
FT	Cleavage-site	1404
FT	/note=	"GPI anchor cleavage site"
XX		
PN	WO200285942-A2.	
XX		
PD	31-OCT-2002.	
XX		
PF	24-APR-2002; 2002WO-CA00560.	
XX		
PR	24-APR-2001; 2001US-285713P.	
XX		
PR	14-FEB-2002; 2002US-356163P.	
XX		

PA (UYMC-) UNIV MCGILL.  
 XX Philip A, Tam B;  
 PI WPI, 2003-093100/08.  
 DR N-PSDB; AAD49434.  
 XX Novel transforming growth factor (TGF)-beta 1 binding reagent which  
 PT comprises r150 protein which acts as accessory receptor of TGF-beta,  
 PT useful for negatively modulating TGF-beta activity, and thus for  
 PT treating cancer -  
 XX  
 PS Claim 1; Page 94-99; 127pp; English.  
 XX  
 CC The invention relates to novel transforming growth factor (TGF)-beta1  
 CC binding reagent which comprises a glycosylphosphatidylinositol (GPI)-  
 CC anchored TGF-beta1 binding protein referred to as r150 which acts as  
 CC an accessory receptor of TGF-beta. The invention is used for negatively  
 CC modulating TGF-beta activity, and thus for treating conditions  
 CC characterised by overproduction of TGF-beta, such as cancer. Antisense  
 CC molecules of the invention are used for increasing TGF-beta availability  
 CC and increase graft success. The present sequence is human r150 protein.  
 XX  
 SQ Sequence 1428 AA;  
 Query Match 99.8%; Score 7331; DB 24; Length 1428;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 1425; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 661 AERFMEBNEGHIVDIDHDFSLGSSPHVRKHPETWIMLDITNMGYRIQEFEEVTVPSDITSW 720  
 Db 661 AERFMEBNEGHIVDIDHDFSLGSSPHVRKHPETWIMLDITNMGYRIQEFEEVTVPSDITSW 720  
 QY 721 VATEGVISEDGLGLTTTPVELQAFPPFIFLNLPSVIRGEFALEITFNILKATEV 780  
 Db 721 VATEGVISEDGLGLTTTPVELQAFPPFIFLNLPSVIRGEFALEITFNILKATEV 780  
 QY 781 KVIIEKSDKFDIMTSEINATGHOOTLLVPSBDGATVFPPIPTHGEPITVTALSP 840  
 Db 781 KVIIEKSDKFDIMTSEINATGHOOTLLVPSBDGATVFPPIPTHGEPITVTALSP 840  
 QY 841 ASDAVTOMILVKAEGIEKSYQSILLDLTDNRLOSTLKLTSFSPPNVTGSSRVOITAI 900  
 Db 841 ASDAVTOMILVKAEGIEKSYQSILLDLTDNRLOSTLKLTSFSPPNVTGSSRVOITAI 900  
 QY 901 GDTLGSINGLALIMPGCCGQNMNINPAPNIYIIDYLTKKQOLDNLKELSPKROG 960  
 Db 901 GDTLGSINGLALIMPGCCGQNMNINPAPNIYIIDYLTKKQOLDNLKELSPKROG 960  
 QY 961 YORELLYOREDGSFSAFGNVPDGSITWLSAFVLRCELEADPYIDIDONVLRHTYWLK 1020  
 Db 961 YORELLYOREDGSFSAFGNVPDGSITWLSAFVLRCELEADPYIDIDONVLRHTYWLK 1020  
 QY 1021 QKNGEFMDPGRVIHSELQGNKSPVTLTAYIVTSLLGYRKYOPNIDVOESIHFLESEFS 1080  
 Db 1021 QKNGEFMDPGRVIHSELQGNKSPVTLTAYIVTSLLGYRKYOPNIDVOESIHFLESEFS 1080  
 QY 1081 RGISDNYTTLALITYALSSVGSPPAKKALNMLTWRAOEGQMOWWSESKLSNQPRSL 1140  
 Db 1081 RGISDNYTTLALITYALSSVGSPPAKKALNMLTWRAOEGQMOWWSESKLSNQPRSL 1140  
 QY 1141 DIEVAAYALSHFLQOTSSEGIPIKMWLSRORSLGFPASTODTPTAALKALSEFALMNT 1200  
 Db 1141 DIEVAAYALSHFLQOTSSEGIPIKMWLSRORSLGFPASTODTPTAALKALSEFALMNT 1200  
 QY 1201 ERTNIGVTVSPSPPLAVVQMAVNIISANGFPAICOLNVVYVYKASGSSRRRSION 1260  
 Db 1201 ERTNIGVTVSPSPPLAVVQMAVNIISANGFPAICOLNVVYVYKASGSSRRRSION 1260  
 QY 1261 QEAFDIDVAVKENKDLNHYDLNVCISFSGRGSMALEVNLLSGFMPSEALISSETV 1320  
 Db 1261 QEAFDIDVAVKENKDLNHYDLNVCISFSGRGSMALEVNLLSGFMPSEALISSETV 1320  
 QY 1321 KKYEDHGKLNILDSVNETQFCVNI PAVANFVSTODASVSIIVYYPEPRQAVASYSN 1380  
 Db 1321 KKYEDHGKLNILDSVNETQFCVNI PAVANFVSTODASVSIIVYYPEPRQAVASYSN 1380  
 QY 1381 EYKLSSCDLCSVDQGRPCEDGASGSHHSSVIFICFKLLYFMEI LML 1428  
 Db 1381 EYKLSSCDLCSVDQGRPCEDGASGSHHSSVIFICFKLLYFMEI LML 1428

RESULT 2  
 ABB82165  
 ID ABB82165 standard; Protein; 1445 AA.  
 AC ABB82165;  
 DT 23-DEC-2002 (first entry)  
 XX  
 DE Human CD109 K1 protein.  
 XX  
 KW CD109; alpha2 macroglobulin; thioester; cerebroprotective; cardiant;  
 KW immunosuppressive; haemostatic; anticoagulant; thrombolytic; human;  
 KW cardiovascular; vasotrophic; gene therapy; CD109 K1.  
 OS Homo sapiens.  
 XX  
 PN W0200270696-A2.  
 XX  
 PD 12-SEP-2002.  
 XX

PF 07-MAR-2002; 2002MO-CA00292.  
 XX 07-MAR-2001; 2001US-273814P.  
 XX (SCHU/) SCHUH A.  
 PA (SUTH/) SUTHERLAND R D.  
 XX Schuh A, Sutherland RD;  
 PI MPI; 2002-713450/77.  
 DR N-PSDB; ABQ79964.  
 XX  
 PT New CD109 nucleic acids and polypeptides, useful in gene therapy,  
 PT particularly for treating strokes, myocardial infarctions, thrombosis,  
 PT thrombocytopenia, autoimmune diseases, or organ or bone marrow  
 PT transplantation -  
 XX  
 PS Claim 21; Fig 3a; 156pp; English.  
 XX  
 CC The invention relates to isolated nucleic acid molecules encoding CD109  
 CC polypeptides. These nucleic acid molecules include the human CDNA  
 CC sequences comprising CD109 K1, CD109 K1-H7, CD109 K15 or their variants.  
 CC CD109 is a novel member of the alpha2 macroglobulin (alpha2M/C3, C4, C5  
 CC family of thioester-containing proteins. The CD109 polypeptides can be  
 CC expressed by standard recombinant methodology. The CD109 nucleic acid,  
 CC CD109 polypeptide or its mimetic, or the anti-CD109 antibody is useful  
 CC for treating or detecting a disease or disorder, e.g. conditions  
 CC associated with endothelial activation, platelet activation, activation  
 CC of the coagulation or fibrinolytic systems, activation of T lymphocytes  
 CC and of the complement system, quantitative or qualitative abnormalities  
 CC of platelet function, increased or impaired platelet aggregation and  
 CC activation, increased or impaired activation of the coagulation and/or  
 CC fibrinolytic systems, or impaired or increased immune activation. These  
 CC are also useful for treating cardiovascular disorders, stroke, myocardial  
 CC infarction, thrombosis, embolism, peripheral vascular disease,  
 CC thrombocytopenia, thrombocythemia, autoimmune diseases, organ  
 CC transplantation, or bone marrow transplantation. The CD109 nucleic acids  
 CC or their antisense nucleotide sequence are useful in gene therapy for  
 CC treating these conditions. The present sequence represents the human  
 CC CD109 K1 sequence.  
 XX  
 SQ Sequence 1445 AA;  
 Query Match 99.7%; Score 7326.5; DB 23; Length 1445;  
 Best Local Similarity 98.8%; Pred. No. 0;  
 Matches 1427; Conservative 1; Mismatches 0; Indels 17; Gaps 1;

DB 361 KPSLNFATVKTADGNQLTEERRNNVITVTQRYNTEYWSGNSGNQKMEAVOKINY 420  
 QY 421 TVPQSGTFKIEPPIEDSSELOLKAYFLGSKSMVAHSLFKSPSKTYIQLKTRDENIKVG 480  
 DB 421 TVPQSGTFKIEPPIEDSSELOLKAYFLGSKSMVAHSLFKSPSKTYIQLKTRDENIKVG 480  
 QY 481 SPFELVSGNKKLKELSYMWVSRGQLVAVGKNSTMFSLTIPENSWTPKACVIVYIIEDDG 540  
 DB 481 SPFELVSGNKKLKELSYMWVSRGQLVAVGKNSTMFSLTIPENSWTPKACVIVYIIEDDG 540  
 QY 541 EIISDVLLKIPVQVVRKNIKLYMSKYKAPSPSKVSLRISVTPDSTVIGIVANDKSVNLNN 600  
 DB 541 EIISDVLLKIPVQVVRKNIKLYMSKYKAPSPSKVSLRISVTPDSTVIGIVANDKSVNLNN 600  
 QY 541 EIISDVLLKIPVQVVRKNIKLYMSKYKAPSPSKVSLRISVTPDSTVIGIVANDKSVNLNN 600  
 DB 601 ASNDITMENVVHELELYNTGYLLGEMNNSFAVFOCGMLVLDANLTKOYIGGVYDNAMEY 660  
 QY 601 ASNDITMENVVHELELYNTGYLLGEMNNSFAVFOCGMLVLDANLTKOYIGGVYDNAMEY 660  
 DB 601 ASNDITMENVVHELELYNTGYLLGEMNNSFAVFOCGMLVLDANLTKOYIGGVYDNAMEY 660  
 QY 661 AERFMEENEGHIVDIDHPSLGSSPHVRKHPETWTWLDTNMGYRIYQEEVTVPSITSM 720  
 DB 661 AERFMEENEGHIVDIDHPSLGSSPHVRKHPETWTWLDTNMGYRIYQEEVTVPSITSM 720  
 QY 721 VATGVVISDGLGLTTTPVELQAPPPFIPLNLPYSYIRGEFALETIYNYLQDATEV 780  
 DB 721 VATGVVISDGLGLTTTPVELQAPPPFIPLNLPYSYIRGEFALETIYNYLQDATEV 780  
 QY 781 KTIIEKSKFDILMTSSRNATGHQOTLLVSEDAVATLPRPHLGEIPITVVALSPT 840  
 DB 781 KTIIEKSKFDILMTSSRNATGHQOTLLVSEDAVATLPRPHLGEIPITVVALSPT 840  
 QY 841 ASDAVTOMILVNAEGIEKYSQSILDLTDNRLOSTLXTLSFPFPNTVTSERVOITAI 900  
 DB 841 ASDAVTOMILVNAEGIEKYSQSILDLTDNRLOSTLXTLSFPFPNTVTSERVOITAI 900  
 QY 901 GVLTPSINGLASLIRMPYCGEQMNPAPNIYILDYTKKKQUTDLKEXALSPMROG 960  
 DB 901 GVLTPSINGLASLIRMPYCGEQMNPAPNIYILDYTKKKQUTDLKEXALSPMROG 960  
 QY 961 YORELLYOREDSFSAFGNYPDSSGTWLSAFLRCLFELADPIDIDQVLAHTYTWLKGH 1020  
 DB 961 YORELLYOREDSFSAFGNYPDSSGTWLSAFLRCLFELADPIDIDQVLAHTYTWLKGH 1020  
 QY 1021 QKSNGEFMDPGRVHISELQGNKSPVTLTAYIVTSLGGRKYQPNIDVQESIHFLSEFS 1080  
 DB 1021 QKSNGEFMDPGRVHISELQGNKSPVTLTAYIVTSLGGRKYQPNIDVQESIHFLSEFS 1080  
 QY 1081 RGISDNYTLALITVALSSVSGSPKAKEALNMLTWRAEOGGMQFWVSSSKLSDSQPRSL 1140  
 DB 1081 RGISDNYTLALITVALSSVSGSPKAKEALNMLTWRAEOGGMQFWVSSSKLSDSQPRSL 1140  
 QY 1141 DIEVAAYALLSHFLOPOTSEGIPIRWLSRQNSLSGFSASTODTTVALKALSEPALNMT 1200  
 DB 1141 DIEVAAYALLSHFLOPOTSEGIPIRWLSRQNSLSGFSASTODTTVALKALSEPALNMT 1200  
 QY 1201 ERTNIQVTVGPSRSP-----LAVQPMANVTSANGFGAICOLNV 1243  
 DB 1201 ERTNIQVTVGPSRSP-----LAVQPMANVTSANGFGAICOLNV 1243  
 QY 1244 YNVKASGSSRRRRRSIONOEAFLDVAVKNKODLNHVDLVNCTSPSGRSGMALMEVNL 1303  
 DB 1244 YNVKASGSSRRRRRSIONOEAFLDVAVKNKODLNHVDLVNCTSPSGRSGMALMEVNL 1303  
 QY 1304 LSGFVVPSEALISLSTVKKVEYDHGKLNLYLDSVNETOPCVNI PAVRNFKVSNTO DASVS 1363  
 DB 1304 LSGFVVPSEALISLSTVKKVEYDHGKLNLYLDSVNETOPCVNI PAVRNFKVSNTO DASVS 1363  
 QY 1364 IVDYIEPRROAVRSVNSSEYKLSCDLCSDVQCGRCCEGASGSHHSVYIFCFKRLYF 1423  
 DB 1364 IVDYIEPRROAVRSVNSSEYKLSCDLCSDVQCGRCCEGASGSHHSVYIFCFKRLYF 1423  
 QY 1424 MELWL 1428  
 DB 1424 MELWL 1445

RESULT 3  
 ID ABB82167 standard; Protein: 1445 AA.  
 AC ABB82167;  
 DT 23-DEC-2002 (first entry)  
 DE Human CD109 KI-H7 protein.  
 OS Homo sapiens.  
 PN MO200270696-A2.  
 PD 12-SEP-2002.  
 PF 07-MAR-2002; 2002MO-CA00292.  
 PR 07-MAR-2001; 2001US-273814P.  
 PA (SCHU/) SCHUH A.  
 PA (SUTH/) SUTHERLAND R D.  
 PI Schuh A, Sutherland RD;  
 DR MPI; 2002-713450/77.  
 DR N-PSDB; ABQ79966.  
 XX  
 PT New CD109 nucleic acid and polypeptides, useful in gene therapy,  
 PT particularly for treating strokes, myocardial infarctions, thrombosis,  
 PT thrombocytopenia, autoimmune diseases, or organ or bone marrow  
 PT transplantation -  
 XX  
 PS Claim 21; Fig 3a; 156pp; English.  
 XX  
 CC The invention relates to isolated nucleic acid molecules encoding CD109  
 CC polypeptides. These nucleic acid molecules include the human cDNA  
 CC sequences comprising CD109 KI, CD109 KI-H7, CD109 K15 or their variants.  
 CC CD109 is a novel member of the alpha2 macroglobulin (alpha2M)/C3, C4, C5  
 CC family of thioester-containing proteins. The CD109 polypeptides can be  
 CC expressed by standard recombinant methodology. The CD109 nucleic acid,  
 CC CD109 polypeptide or its mimetic, or the anti-CD109 antibody is useful  
 CC for treating or detecting a disease or disorder, e.g. conditions  
 CC associated with endothelial activation, platelet activation, activation  
 CC of the coagulation or fibrinolytic systems, activation of T lymphocytes  
 CC and of the complement system, quantitative or qualitative abnormalities  
 CC of platelet function, increased or impaired platelet aggregation and  
 CC activation, increased or impaired activation of the coagulation and/or  
 CC fibrinolytic systems, or impaired or increased immune activation. These  
 CC are also useful for treating cardiovascular disorders, stroke, myocardial  
 CC infarction, thrombosis, embolism, peripheral vascular disease,  
 CC thrombocytopenia, thrombocytopenia, autoimmune diseases, organ  
 CC transplantation, or bone marrow transplantation. The CD109 nucleic acids  
 CC or their antisense nucleotide sequence are useful in gene therapy for  
 CC treating these conditions. The present sequence represents the human  
 CC CD109 KI-H7 sequence.  
 XX  
 SQ Sequence 1445 AA;  
 Query Match 99.7%; Score 7326.5; DB 23; Length 1445;  
 Best Local Similarity 98.8%; Pred. No. 0;  
 Matches 1427; Conservative 1; Mismatches 0; Indels 17; Gaps 1;  
 QY 1 MOSPRLTAAHLTCVCTAALAVAGPRFLVTAGPITPPGNVITGVLLHCHSQTIVKA 60  
 DB 1 MOSPRLTAAHLTCVCTAALAVAGPRFLVTAGPITPPGNVITGVLLHCHSQTIVKA 60

QY 61 ELIKTASNLTVSLVLAEGVFEEKSPKTLTLPGLPLNSADEIYELAVTGTODEILPSNST 120  
 DB 61 ELIKTASNLTVSLVLAEGVFEEKSPKTLTLPGLPLNSADEIYELAVTGTODEILPSNST 120  
 QY 121 RLSFETKRISVFTQTKALYKPKQEVKFRIVLTFSPFKPKTSINTLIDPKSNLQQWL 180  
 DB 121 RLSFETKRISVFTQTKALYKPKQEVKFRIVLTFSPFKPKTSINTLIDPKSNLQQWL 180  
 QY 181 SQQSDLGIVISKTFQLSHPILGDMSIQOVNDQTYQSFQVSEYVLPKEVTLQTFLYCS 240  
 DB 181 SQQSDLGIVISKTFQLSHPILGDMSIQOVNDQTYQSFQVSEYVLPKEVTLQTFLYCS 240  
 QY 241 MNSKHLNGITTAKYTGKPKVGDVTLPLPLSFWSKKKNTTKFKINGSANFSFNDENK 300  
 DB 241 MNSKHLNGITTAKYTGKPKVGDVTLPLPLSFWSKKKNTTKFKINGSANFSFNDENK 300  
 QY 301 NMDSSNGLSSEYLDSSPGVEILTTVTSSVNGISNNVSTNVEFKQHDYIEFDDYTVL 360  
 DB 301 NMDSSNGLSSEYLDSSPGVEILTTVTSSVNGISNNVSTNVEFKQHDYIEFDDYTVL 360  
 QY 361 KPSLNTFAATVKTTRADGNQTLTEERRNNVITYTQRTNYEYWSGNSGNQKMAVOKINY 420  
 DB 361 KPSLNTFAATVKTTRADGNQTLTEERRNNVITYTQRTNYEYWSGNSGNQKMAVOKINY 420  
 QY 421 TVPQSGTFKIEPPIEDSSELQKAYFLGSKSMAVHSLFKSPSKTYIQLTREDNIRVG 480  
 DB 421 TVPQSGTFKIEPPIEDSSELQKAYFLGSKSMAVHSLFKSPSKTYIQLTREDNIRVG 480  
 QY 481 SPFELVSGNKRKLELSYMWVSRGQIVANGKONSTFSLTPENSWTPKACVITYYIEDDG 540  
 DB 481 SPFELVSGNKRKLELSYMWVSRGQIVANGKONSTFSLTPENSWTPKACVITYYIEDDG 540  
 QY 541 EISIDVLKIPVOLVFNKIKLYMSKYKAESEKVSIRISVTQPSIVGIVADVKSANLNM 600  
 DB 541 EISIDVLKIPVOLVFNKIKLYMSKYKAESEKVSIRISVTQPSIVGIVADVKSANLNM 600  
 QY 601 ASNDITMENVHLELNTGYIYGMFNSPAVQECGLWVLDANITKQYIDGVYDIAEY 660  
 DB 601 ASNDITMENVHLELNTGYIYGMFNSPAVQECGLWVLDANITKQYIDGVYDIAEY 660  
 QY 661 AERFMENEGHIVDIDHDFSLGSSPHYRKHPPEWIMLDNNMGRIYOEPRVYVPSDITSW 720  
 DB 661 AERFMENEGHIVDIDHDFSLGSSPHYRKHPPEWIMLDNNMGRIYOEPRVYVPSDITSW 720  
 QY 721 VANGFVISEDGLGLTTTPVELQAFQPFILNLPYSVIRGEFALERTFNLYKQATEV 780  
 DB 721 VANGFVISEDGLGLTTTPVELQAFQPFILNLPYSVIRGEFALERTFNLYKQATEV 780  
 QY 781 KVIEKSDKFDILMTSSEINATGHOQTLVPSDEGATVLPPIRPHLGEIPITVTLSP 840  
 DB 781 KVIEKSDKFDILMTSSEINATGHOQTLVPSDEGATVLPPIRPHLGEIPITVTLSP 840  
 QY 841 ASDAVTOMILVAKEGIEKYSQSILDLTDNRLOSTLKTLSRFPNTYGSRRVQITAI 900  
 DB 841 ASDAVTOMILVAKEGIEKYSQSILDLTDNRLOSTLKTLSRFPNTYGSRRVQITAI 900  
 QY 901 GDTVAGSINGLASLIMPAGCGEONMIFAPNITYIDYVTKKKQLDNLKEKLSFMRQ 960  
 DB 901 GDTVAGSINGLASLIMPAGCGEONMIFAPNITYIDYVTKKKQLDNLKEKLSFMRQ 960  
 QY 961 YORELLYOREBDSFSAFGNYDPSGTSWLSAFVLRCELEADPYIIDQNVLAHRTYWLKG 1020  
 DB 961 YORELLYOREBDSFSAFGNYDPSGTSWLSAFVLRCELEADPYIIDQNVLAHRTYWLKG 1020  
 QY 1021 QKNGEFPMDPGRVYHSELQGNKSPVTLTAIYVTSILGRTKQPNIDVQBSHFLESF 1080  
 DB 1021 QKNGEFPMDPGRVYHSELQGNKSPVTLTAIYVTSILGRTKQPNIDVQBSHFLESF 1080  
 QY 1081 RGISDNYTTLATLYALSSVSPKAKALNMLTWRABEGGOMQWVSSKSLDSMQPRSL 1140  
 DB 1081 RGISDNYTTLATLYALSSVSPKAKALNMLTWRABEGGOMQWVSSKSLDSMQPRSL 1140  
 QY 1141 DIEVAAYALLSHFLQFOTSEGIPIMRWLSRORNSLGGFASOTDTVALKALSEFALMNT 1200

```

Db      1141 DIEVAAYALLSHFLQFOTISEGIPIMRWLSRORNSJGGFASITODTTVALKAUSEFALMNT 1200
Qy      1201 ERTNIQVTVTPSPSPSP-----LAVVQPMANISANGFGAICQLNV 1243
Db      1201 ERTNIQVTVTPSPSPSPVKFLIDITHNRLLOTAELAVVQPMANISANGFGAICQLNV 1260
Qy      1244 YNVKASGSSRRRRRSIQNOBAPDLVAVKENKODLNHVDLNVCTSFSGPGRSGMALMEVNL 1303
Db      1261 YNVKASGSSRRRRRSIQNOBAPDLVAVKENKODLNHVDLNVCTSFSGPGRSGMALMEVNL 1320
Qy      1304 LSGFVWPEBASTLSTETVKKVEVDHCKLNLXLDVSVNETQCVNIPAVRNPKVSNTOBASYS 1363
Db      1321 LSGFVWPEBASTLSTETVKKVEVDHCKLNLXLDVSVNETQCVNIPAVRNPKVSNTOBASYS 1380
Qy      1364 IVDYEPFRQAVRSYNSEVKLSSCDLSCDVQGCRCPCEDGASGSHHSVTFIFCFKLLYF 1423
Db      1381 IVDYEPFRQAVRSYNSEVKLSSCDLSCDVQGCRCPCEDGASGSHHSVTFIFCFKLLYF 1440
Qy      1424 MELWL 1428
Db      1441 MELWL 1445

```

RESULT 4  
AA019372 standard; Protein; 1445 AA.

AA019372;

04-DEC-2002 (first entry)

Human platelet alloantigen Gova.

Human; platelet alloantigen; Gova; Govb; single nucleotide polymorphism;

SNP; biallelic; bleeding disorder; post-transfusion purpura;

post-transfusion platelet refractoriness; haemostatic; vaccine;

neonatal alloimmune thrombocytopenia.

Homo sapiens.

MO200270738-A2.

12-SEP-2002.

07-MAR-2002; 2002MO-CA00291.

07-MAR-2001; 2001US-273941P.

(SCHU/) SCHUH A.

Schuh A, Ouwehand W;

WPI; 2002-713460/77.

N-PSDB; AAL49815.

New isolated oligonucleotide binding to a region of CD109 nucleic acid

having a single nucleotide polymorphism that distinguishes a Gova

and/or Govb allele, useful for treating blood disorders e.g. alloimmune

thrombocytopenia

Disclosure; Page 29-35; 69pp; English.

The present invention relates to a sequence capable of binding

specifically to a CD109 nucleic acid which has a single nucleotide

polymorphism that distinguishes the Gova and Govb alleles. Detection of

the Gov genotype is useful for detecting whether the subject has or is at

risk of a blood disease, disorder or abnormal physical state, such as

bleeding, or increased risk of bleeding, due to alloimmune destruction of

blood platelets, e.g., post-transfusion purpura, post-transfusion

platelet refractoriness or neonatal alloimmune thrombocytopenia. The

nucleic acid and polypeptide are useful for Gov genotyping or phenotyping

individuals. The present sequence is the Gova encoded protein.

XX Sequence 1445 AA;  
SQ  
Query Match 99.7%; Score 7326.5; DB 23; Length 1445;  
Best Local Similarity 98.8%; Pred No. 0;  
Matches 1427; Conservative 1; Mismatches 0; Indels 17; Gaps 1;

```

Qy      1 MOGPBLTAAHLCTCTAALAVAPGPRLVTAAGIIRPGNVITGVLELHCPQVATKA 60
Db      1 MOGPBLTAAHLCTCTAALAVAPGPRLVTAAGIIRPGNVITGVLELHCPQVATKA 60
Qy      61 ELTKASNLTVSVLEAEVFEKSGFKTLTLPPLNSADEIYEALRVGTODEILFNSNT 120
Db      61 ELTKASNLTVSVLEAEVFEKSGFKTLTLPPLNSADEIYEALRVGTODEILFNSNT 120
Qy      121 RLSEFTKRSVFIQTDKALYKPKQVKKRYITLTSDFPKYTSINILIKDKPSNLIQWL 180
Db      121 RLSEFTKRSVFIQTDKALYKPKQVKKRYITLTSDFPKYTSINILIKDKPSNLIQWL 180
Qy      181 SQQSDLAGVISKTFQLSHPILGWSIQVQVNDQTYQSFQSEVYLPRFEVTLQTPLYCS 240
Db      181 SQQSDLAGVISKTFQLSHPILGWSIQVQVNDQTYQSFQSEVYLPRFEVTLQTPLYCS 240
Qy      241 MNSKRLNGTITAKTYGKPKVGDVTLTFLPLSPFGKKKNTTTPKINGSANFSFDEBKM 300
Db      241 MNSKRLNGTITAKTYGKPKVGDVTLTFLPLSPFGKKKNTTTPKINGSANFSFDEBKM 300
Qy      301 NVMDSSNGLSEYLDLSSGPVEILTTVRESVTGISRNVSTNVFQKHYYIIIEFDYTVL 360
Db      301 NVMDSSNGLSEYLDLSSGPVEILTTVRESVTGISRNVSTNVFQKHYYIIIEFDYTVL 360
Qy      361 KPSLNFATVATVTAADGNQLTLERRNNVITVQRYNTEYWSGSGNSGNQKMEAVOKINY 420
Db      361 KPSLNFATVATVTAADGNQLTLERRNNVITVQRYNTEYWSGSGNSGNQKMEAVOKINY 420
Qy      421 TVPQSGTFKIEPPILEDSSSELQKAYFLGSKSSMAVHSLFKSPKTYIQLKTRDENIKYG 480
Db      421 TVPQSGTFKIEPPILEDSSSELQKAYFLGSKSSMAVHSLFKSPKTYIQLKTRDENIKYG 480
Qy      481 SPFELVSGNKRLELSYVVSRCQLVAVGQNSTMTSLTPENSGTPACVAYVYIEDDG 540
Db      481 SPFELVSGNKRLELSYVVSRCQLVAVGQNSTMTSLTPENSGTPACVAYVYIEDDG 540
Qy      541 EISDVLKIPVQVLPKNIKLYMSKVKAPESEKSLRSVQPSIVGIVAVDSVNLNN 600
Db      541 EISDVLKIPVQVLPKNIKLYMSKVKAPESEKSLRSVQPSIVGIVAVDSVNLNN 600
Qy      601 ASNDITMENVHLELYNTGYLLGMFNNSFAVFOECGLMVLTDANLTQDYIDGVYDNAEY 660
Db      601 ASNDITMENVHLELYNTGYLLGMFNNSFAVFOECGLMVLTDANLTQDYIDGVYDNAEY 660
Qy      661 AERFMEENEGHIVDIHDFSLGSSPHVRKHPERTIMWLDTNMGYRIYQFEVTVPDSITSW 720
Db      661 AERFMEENEGHIVDIHDFSLGSSPHVRKHPERTIMWLDTNMGYRIYQFEVTVPDSITSW 720
Qy      721 VATGFVISEDGLGTTTPVELQAPQPFITLNPYSYIRGEFPLETITNNYLKDADEV 780
Db      721 VATGFVISEDGLGTTTPVELQAPQPFITLNPYSYIRGEFPLETITNNYLKDADEV 780
Qy      781 KVILKSDKPDILMTSSSEINATGHQOTLVSESDGATVLPFIRPTHGAEIPIYVATASPT 840
Db      781 KVILKSDKPDILMTSSSEINATGHQOTLVSESDGATVLPFIRPTHGAEIPIYVATASPT 840
Qy      841 ASDAVTQMITLVKAGIEKSYQSILDLTDNRLOSTLKTLSFSPPNVTVSERVQITAI 900
Db      841 ASDAVTQMITLVKAGIEKSYQSILDLTDNRLOSTLKTLSFSPPNVTVSERVQITAI 900
Qy      901 GDVIGPSINGLASLIRMPYGGGEONMNFANITYILDTYTKKQLTDLKKAISPFMKG 960
Db      901 GDVIGPSINGLASLIRMPYGGGEONMNFANITYILDTYTKKQLTDLKKAISPFMKG 960
Qy      961 YQRELLYQREDGSFAFGNYPDSGTWLSAVELRCFLEADPYIDQNVLHRTYTWLKGH 1020
Db      961 YQRELLYQREDGSFAFGNYPDSGTWLSAVELRCFLEADPYIDQNVLHRTYTWLKGH 1020

```

Db 961 YQRELLYQREDEGFSFAGNYPDPSGTLWSAFVLCFLEADPYIDIDONVLRHYTWLKGH 1020  
QY 1021 QKNGEFMDPGRVHSELQGNKSPVTLTAIVTSLGKYKQPNIDVOESIHFLSESEFS 1080  
XX 1021 QKNGEFMDPGRVHSELQGNKSPVTLTAIVTSLGKYKQPNIDVOESIHFLSESEFS 1080  
Db 1021 QKNGEFMDPGRVHSELQGNKSPVTLTAIVTSLGKYKQPNIDVOESIHFLSESEFS 1080  
QY 1081 RGISDNTTALITYALSSVSGSPKAKELNMLTWBAEBOGQMFVWSESEKLSDSWQPSRL 1140  
Db 1081 RGISDNTTALITYALSSVSGSPKAKELNMLTWBAEBOGQMFVWSESEKLSDSWQPSRL 1140  
QY 1141 DIEVAALSHPLQOPTSGIPIMRWLSRORNSLGFPASTQDTTVALKALSEFALAMNT 1200  
Db 1141 DIEVAALSHPLQOPTSGIPIMRWLSRORNSLGFPASTQDTTVALKALSEFALAMNT 1200  
QY 1201 ERTNIQVTVTGPSSPSP-----LAVQPMVANI SANGFPAICQLNVV 1243  
Db 1201 ERTNIQVTVTGPSSPSP-----LAVQPMVANI SANGFPAICQLNVV 1243  
QY 1244 YNVKASGSSRRRSIQNOEAFDLDAVAKENKDLNHDVLANVCTSGSPGRSGMALMEVNL 1303  
Db 1261 YNVKASGSSRRRSIQNOEAFDLDAVAKENKDLNHDVLANVCTSGSPGRSGMALMEVNL 1320  
QY 1304 LSGFMVSEALSLSEYTKVKEVDHGLNLYDSVNEFOFCVNIIPAVNPFKVSNTQDASVS 1363  
Db 1321 LSGFMVSEALSLSEYTKVKEVDHGLNLYDSVNEFOFCVNIIPAVNPFKVSNTQDASVS 1380  
QY 1364 IVDYEBRROAVRSYNESEVKLSGDLCSDVOCGRPCEDGASGSHHSSVIFIFCFKLLYF 1423  
Db 1381 IVDYEBRROAVRSYNESEVKLSGDLCSDVOCGRPCEDGASGSHHSSVIFIFCFKLLYF 1440  
QY 1424 MELWL 1428  
Db 1441 MELWL 1445

## RESULT 5

AAE32013 ID AAE32013 standard; Protein; 1445 AA.

XX AAE32013;

AC AAE32013;

DT 24-MAR-2003 (first entry)

DE Human r150 protein #2.

XX Human: GpI-anchored TGF-beta1 binding protein; r150 protein; cancer;

KW glycoylphosphatidyl inositol; transforming growth factor-beta1;

KM therapy.

XX Homo sapiens.

OS Homo sapiens.

XX Key

FT Misc-difference

FT /label= Ser, Tyr

XX /note= "Encoded by TNC"

XX WO200265942-A2.

XX 31-OCT-2002.

XX 24-APR-2002; 2002WO-CA00560.

XX 24-APR-2001; 2001US-285713P.

XX 14-FEB-2002; 2002US-356163P.

XX (UYMC-) UNIV MCGIL.

XX Philip A, Tam B;

XX WPI, 2003-093100/08.

XX N-PSDB; AAD49435.

XX Novel transforming growth factor (TGF)-beta 1 binding reagent which

PT comprises r150 protein which acts as accessory receptor of TGF-beta,  
PT useful for negatively modulating TGF-beta activity, and thus for  
PT treating cancer -  
PS Claim 1; Page 103-109; 127pp; English.  
XX The invention relates to novel transforming growth factor (TGF)-beta1  
CC binding reagent which comprises a glycosylphosphatidyl inositol (GPI)-  
CC anchored TGF-beta1 binding protein referred to as r150 which acts as  
CC an accessory receptor of TGF-beta. The invention is used for negatively  
CC modulating TGF-beta activity, and thus for treating conditions  
CC characterised by overproduction of TGF-beta, such as cancer. Antisense  
CC molecules of the invention are used for increasing TGF-beta availability  
CC and increase graft success. The present sequence is human r150 protein.  
XX  
SQ Sequence 1445 AA;

Query Match 99.6%; Score 7318.5; DB 24; Length 1445;

Best Local Similarity 98.7%; Pred. No. 0;

Matches 1426; Conservative 1; Mismatches 1; Indels 17; Gaps 1;

QY 1 MOGPPLTAAHLTCVTAALAVAPGRPLVTARGIIRPGSNVTIGVELHCHPSQVTVKA 60  
Db 1 MOGPPLTAAHLTCVTAALAVAPGRPLVTARGIIRPGSNVTIGVELHCHPSQVTVKA 60  
QY 61 ELIKTASNLTVSLAEAGVPEKESFKTLTLPSPILNSADEIYELRYGTODEILFSNST 120  
Db 61 ELIKTASNLTVSLAEAGVPEKESFKTLTLPSPILNSADEIYELRYGTODEILFSNST 120  
QY 121 RLSFETKRISVFIQTDKALYKQEVKFRIVTLPSDPKPYKTSNLILKOPKSNLIQOWL 180  
Db 121 RLSFETKRISVFIQTDKALYKQEVKFRIVTLPSDPKPYKTSNLILKOPKSNLIQOWL 180  
QY 181 SQOSDLGVSKTEFQLSHPLTGMSTIOVQNDQTYQSFQVSEYVLPKPEVTLQTPLYCS 240  
Db 181 SQOSDLGVSKTEFQLSHPLTGMSTIOVQNDQTYQSFQVSEYVLPKPEVTLQTPLYCS 240  
QY 241 MNSKHLNGTITAKYTKGPKYKGVDTLTPLPSPFGKKKNTTKPKINGSNANFSNDEMK 300  
Db 241 MNSKHLNGTITAKYTKGPKYKGVDTLTPLPSPFGKKKNTTKPKINGSNANFSNDEMK 300  
QY 301 NWDSSNGSLSEYDLSSPGVEILTTVTBESVTGISRNVSTNVFPKQHDYIIEFDDYTVL 360  
Db 301 NWDSSNGSLSEYDLSSPGVEILTTVTBESVTGISRNVSTNVFPKQHDYIIEFDDYTVL 360  
QY 361 KPSLNFATATKYTRAGNOLTLERRNNVITYTORNYEYWGSGNSGNQKMAVQKINY 420  
Db 361 KPSLNFATATKYTRAGNOLTLERRNNVITYTORNYEYWGSGNSGNQKMAVQKINY 420  
QY 421 TVPQSGTFKIEPILDESSSELQKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVG 480  
Db 421 TVPQSGTFKIEPILDESSSELQKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVG 480  
QY 481 SPFELVSGNKRKLEISYVWSRGQLVAVGKQNSTWESLTPENSWPKACVIYYIEDDG 540  
Db 481 SPFELVSGNKRKLEISYVWSRGQLVAVGKQNSTWESLTPENSWPKACVIYYIEDDG 540  
QY 541 EIIISDVLTIPVOLVFKKKIKLGYMSKYKABSEKVSRLISTQDSDIYGIYAVVKSVMNM 600  
Db 541 EIIISDVLTIPVOLVFKKKIKLGYMSKYKABSEKVSRLISTQDSDIYGIYAVVKSVMNM 600  
QY 601 ASNDITMENVHELELYNTGYLGMFMSPAVFOEGMWLTDPANITKDIYDGYDNAEY 660  
Db 601 ASNDITMENVHELELYNTGYLGMFMSPAVFOEGMWLTDPANITKDIYDGYDNAEY 660  
QY 661 AERFMEENEGHIVDHFSLGSSPHVRKHPETWIMLDTWQXRIYQEFVTVYDSDITSW 720  
Db 661 AERFMEENEGHIVDHFSLGSSPHVRKHPETWIMLDTWQXRIYQEFVTVYDSDITSW 720  
QY 721 VATGFVISEDGLGLTTTPVELQAFQFPFIFNLVPSVIRGSEFALBITFNILKQATEV 780  
Db 721 VATGFVISEDGLGLTTTPVELQAFQFPFIFNLVPSVIRGSEFALBITFNILKQATEV 780



```

QY 781 KXIEKSDKEDILMTSSSEINATGHQOITLVPSFDGATVLPPIRPTHGSIPIITVTLASTP 840
DB 781 KXIEKSDKEDILMTSSSEINATGHQOITLVPSFDGATVLPPIRPTHGSIPIITVTLASTP 840
QY 841 ASDAVTOMLVAAEGIEKSYOSIILDLIDNRLQSTLKTLSFSPFNWTVGSEKRVQITAI 900
DB 841 ASDAVTOMLVAAEGIEKSYOSIILDLIDNRLQSTLKTLSFSPFNWTVGSEKRVQITAI 900
QY 901 GDLVGSINGLASLIRMPYCGEQNNINPAPNIYIIDLITLKKQLTDNLKEKALSMROG 960
DB 901 GDLVGSINGLASLIRMPYCGEQNNINPAPNIYIIDLITLKKQLTDNLKEKALSMROG 960
QY 961 YRELLIYQREDSFSAFGNYDPSGSTWLSAFVLRCLFLEADPYIDIDQNLHRTYTWLKH 1020
DB 961 YRELLIYQREDSFSAFGNYDPSGSTWLSAFVLRCLFLEADPYIDIDQNLHRTYTWLKH 1020
QY 1021 QKSNGEFMDGRVTHSELQGNKSPVTLTAYITSLIGRKYQNPINDVESIHFLSESPS 1080
DB 1021 QKSNGEFMDGRVTHSELQGNKSPVTLTAYITSLIGRKYQNPINDVESIHFLSESPS 1080
QY 1081 RGISDNYTLALITVALSSVSGSPKAKALNMLTWRAEQEGMOPWVSSEKLSDSMQPRSL 1140
DB 1081 RGISDNYTLALITVALSSVSGSPKAKALNMLTWRAEQEGMOPWVSSEKLSDSMQPRSL 1140
QY 1141 DIEVAAYALLSHFLOFQISEGIPIMKWSRQNSLGGFASSTQDTVALKALSEFALNMT 1200
DB 1141 DIEVAAYALLSHFLOFQISEGIPIMKWSRQNSLGGFASSTQDTVALKALSEFALNMT 1200
QY 1201 ERTNIQVTTYGSSPSPP-----LAVQPMANISANGPFAICQLANV 1243
DB 1201 ERTNIQVTTYGSSPSPPKFLIDTHNRLLIQTAELAVQPMANISANGPFAICQLANV 1260
QY 1244 YVVKASGSSRRRSRISNOCAFILDVAVKENKDLINHDVNLVCTSFSGPGRSGMALMEVNL 1303
DB 1244 YVVKASGSSRRRSRISNOCAFILDVAVKENKDLINHDVNLVCTSFSGPGRSGMALMEVNL 1303
QY 1261 YVVKASGSSRRRSRISNOCAFILDVAVKENKDLINHDVNLVCTSFSGPGRSGMALMEVNL 1320
DB 1261 YVVKASGSSRRRSRISNOCAFILDVAVKENKDLINHDVNLVCTSFSGPGRSGMALMEVNL 1320
QY 1304 LSGFVWPSEAIISLSETVKVEYDHGKLANLYLDSVNETQFCVNI PAVRNFKVSTQDASVS 1363
DB 1304 LSGFVWPSEAIISLSETVKVEYDHGKLANLYLDSVNETQFCVNI PAVRNFKVSTQDASVS 1363
QY 1321 LSGFVWPSEAIISLSETVKVEYDHGKLANLYLDSVNETQFCVNI PAVRNFKVSTQDASVS 1380
DB 1321 LSGFVWPSEAIISLSETVKVEYDHGKLANLYLDSVNETQFCVNI PAVRNFKVSTQDASVS 1380
QY 1364 IVDYIEPRRQAVRSYNSYKLSGCDLCSVQCGRPEDDASGSHHSSVIFPCFGLIYF 1423
DB 1364 IVDYIEPRRQAVRSYNSYKLSGCDLCSVQCGRPEDDASGSHHSSVIFPCFGLIYF 1423
QY 1381 IVDYIEPRRQAVRSYNSYKLSGCDLCSVQCGRPEDDASGSHHSSVIFPCFGLIYF 1440
DB 1381 IVDYIEPRRQAVRSYNSYKLSGCDLCSVQCGRPEDDASGSHHSSVIFPCFGLIYF 1440
QY 1424 MELML 1428
DB 1441 MELML 1445

```

RESULT 6  
ABB82166  
ID ABB82166 standard; Protein; 1445 AA.

AC ABB82166;  
DT 23-DEC-2002 (first entry)  
DE Human CD109 K1 variant protein.  
XX  
XX CD109; alpha2 macroglobulin; thioester; cerebroprotective; cardiant;  
KW immunosuppressive; haemostatic; anticoagulant; thrombolytic; human;  
XX cardiovascular; vasotropic; gene therapy; CD109 K1; variant.  
OS Homo sapiens.  
XX  
XX WO2002/0696-A2.  
PD 12-SEP-2002.  
XX  
XX 07-MAR-2002; 2002WO-CA00292.  
PF  
XX 07-MAR-2001; 2001US-27381AP.  
XX

PA (SCHU/) SCHUH A.  
PA (SUTH/) SUTHERLAND R D.  
XX  
PI Schuh A, Sutherland RD;  
XX  
DR WPI; 2002-713450/77.  
XX N-PSDB; ABQ79965.  
XX  
PT New CD109 nucleic acids and polypeptides, useful in gene therapy,  
PT particularly for treating strokes, myocardial infarctions, thrombosis,  
PT thrombocytopenia, autoimmune diseases, or organ or bone marrow  
PT transplantation -  
XX  
XX Claim 21; Fig 3b; 156pp; English.  
XX  
CC The invention relates to isolated nucleic acid molecules encoding CD109  
CC polypeptides. These nucleic acid molecules include the human cDNA  
CC sequences comprising CD109 K1, CD109 K1-H7, CD109 K1S or their variants.  
CC CD109 is a novel member of the alpha2 macroglobulin (alpha2M)/C3, C4, C5  
CC family of thioester-containing proteins. The CD109 polypeptides can be  
CC expressed by standard recombinant methodology. The CD109 nucleic acid,  
CC CD109 polypeptide or its mimetic, or the anti-CD109 antibody is useful  
CC for treating or detecting a disease or disorder, e.g. conditions  
CC associated with endothelial activation, platelet activation, activation  
CC of the coagulation or fibrinolytic systems, activation of T lymphocytes  
CC and of the complement system, quantitative or qualitative abnormalities  
CC of platelet function, increased or impaired platelet aggregation and  
CC activation, increased or impaired activation of the coagulation and/or  
CC fibrinolytic systems, or impaired or increased immune activation. These  
CC are also useful for treating cardiovascular disorders, stroke, myocardial  
CC infarction, thrombosis, embolism, peripheral vascular disease,  
CC thrombocytopenia, thrombocythemia, autoimmune diseases, organ  
CC transplantation, or bone marrow transplantation. The CD109 nucleic acids  
CC or their antisense nucleotide sequence are useful in gene therapy for  
CC treating these conditions. The present sequence represents the human  
CC CD109 K1 variant sequence.  
XX  
SQ Sequence 1445 AA;  
XX  
Query Match 99.6%; Score 7317.5; DB 23; Length 1445;  
Best Local Similarity 98.7%; Pred. No. 0;  
Matches 1426; Conservative 1; Mismatches 1; Indels 17; Gaps 1;  
XX  
QY 1 MCGPPLTAAHLCTCTAALAVAPGRPLVTPAGIIRPGANTVIGLELHCPQVTVAA 60  
DB 1 MCGPPLTAAHLCTCTAALAVAPGRPLVTPAGIIRPGANTVIGLELHCPQVTVAA 60  
QY 61 ELKKTASNLTVSVLEAEGVFEKSGFKTLTPSLPINSADIEYELRVGTODEILFSNST 120  
DB 61 ELKKTASNLTVSVLEAEGVFEKSGFKTLTPSLPINSADIEYELRVGTODEILFSNST 120  
QY 121 RLSFETKRISVFIQTDKALKYKQVKKRIYVLBSDFPKYKTSNLILIDPKPSNLIQQL 180  
DB 121 RLSFETKRISVFIQTDKALKYKQVKKRIYVLBSDFPKYKTSNLILIDPKPSNLIQQL 180  
QY 181 SQQSDLGVIKTFQSLSHPIIDGMSIQVANDQTYGQVSEYLVKPEVTLQTPLYCS 240  
DB 181 SQQSDLGVIKTFQSLSHPIIDGMSIQVANDQTYGQVSEYLVKPEVTLQTPLYCS 240  
QY 241 MNSKHLNGITITAKTYTGKPVKGDVLTPLPSFMGKKKNIITFKINGSANSPFNDEMK 300  
DB 241 MNSKHLNGITITAKTYTGKPVKGDVLTPLPSFMGKKKNIITFKINGSANSPFNDEMK 300  
QY 301 NMDSNGISEYLDSSPEPVEILTTVYESVTGISRNVSTNVFFQHDYIIIEFDYTVL 360  
DB 301 NMDSNGISEYLDSSPEPVEILTTVYESVTGISRNVSTNVFFQHDYIIIEFDYTVL 360  
QY 361 KPSLNFATVKTTRADGNOLTEERRNNVITVQRTNTEYVSGNSGQKAEAVQKINY 420  
DB 361 KPSLNFATVKTTRADGNOLTEERRNNVITVQRTNTEYVSGNSGQKAEAVQKINY 420  
QY 421 TVPQGTPIEPIILDESSELQKAYFLGSKSMAVHSLFKSPSKTYIQLKTRDENIKVG 480  
DB 421 TVPQGTPIEPIILDESSELQKAYFLGSKSMAVHSLFKSPSKTYIQLKTRDENIKVG 480

```

Db 421 TVPSGTFKIEPFLIEDSSSELQLAAYFLGSSKSNVAHSLFKSPSKTYIQLKTRDENIKVG 480
Qy 481 SPFLVVSNGKRLKELSYMVVSRGOLVAVGKONSTMSLTPENSMTPACVIYYIEDDG 540
Db 481 SPFLVVSNGKRLKELSYMVVSRGOLVAVGKONSTMSLTPENSMTPACVIYYIEDDG 540
Qy 541 EIIISDVLPVQVLPVFNKKIKLYMSKVAEPSEKVSRLISVTPDSIVGIVAVDKSVNLNM 600
Db 541 EIIISDVLPVQVLPVFNKKIKLYMSKVAEPSEKVSRLISVTPDSIVGIVAVDKSVNLNM 600
Qy 601 ASNDITMENVHLELVTGTYGLGMFVNSPAVPOECGMVLTDLNLTQDYIDGYDVAEX 660
Db 601 ASNDITMENVHLELVTGTYGLGMFVNSPAVPOECGMVLTDLNLTQDYIDGYDVAEX 660
Qy 661 AERPMENEGHIVIHDFSLGSSPHVRKHPETIMIDTNNGYRIYOEFEVTVDSITSM 720
Db 661 AERPMENEGHIVIHDFSLGSSPHVRKHPETIMIDTNNGYRIYOEFEVTVDSITSM 720
Qy 721 VATGFVISEDGLGITTVPVLAQFQPPFIFLNPYSYIRGEFPALEITIFNYLKDAATEV 780
Db 721 VATGFVISEDGLGITTVPVLAQFQPPFIFLNPYSYIRGEFPALEITIFNYLKDAATEV 780
Qy 781 KVITEKSDKPIIIMTSEFINATGHOQTLVPSDEGATVLPFIRPTHGELPITVTAISPT 840
Db 781 KVITEKSDKPIIIMTSEFINATGHOQTLVPSDEGATVLPFIRPTHGELPITVTAISPT 840
Qy 841 ASDAVTOMILVKAEGIEKYSQSILDLDTNRLQSTLKTSSFPSPPTVTSSEVQITAI 900
Db 841 ASDAVTOMILVKAEGIEKYSQSILDLDTNRLQSTLKTSSFPSPPTVTSSEVQITAI 900
Qy 901 GDVLPSINGLASLIRPYGCGEONMIFAPNIYILDYLTKKQLTDLNKEKALSPKQ 960
Db 901 GDVLPSINGLASLIRPYGCGEONMIFAPNIYILDYLTKKQLTDLNKEKALSPKQ 960
Qy 961 YQRELLVQREBSSASAGNTPSSSTWLSAFVACPLPADYIDIDQNVLRHTTWLKGH 1020
Db 961 YQRELLVQREBSSASAGNTPSSSTWLSAFVACPLPADYIDIDQNVLRHTTWLKGH 1020
Qy 1021 QKSGEFMDPGRVHSELQGNKSPVTLTAYITVSLGKYKYPINIDVQESIHFLSEFS 1080
Db 1021 QKSGEFMDPGRVHSELQGNKSPVTLTAYITVSLGKYKYPINIDVQESIHFLSEFS 1080
Qy 1081 RGISDNTYLLALITYALS SVGSPKAKKALNMLTWRABEGNQFVWSSESKLSDSQPRSL 1140
Db 1081 RGISDNTYLLALITYALS SVGSPKAKKALNMLTWRABEGNQFVWSSESKLSDSQPRSL 1140
Qy 1141 DIEVAAYALLSHFQFOTSEGIPIRMWLSRQNSLGGFASQDTTVALKALSEPALMNT 1200
Db 1141 DIEVAAYALLSHFQFOTSEGIPIRMWLSRQNSLGGFASQDTTVALKALSEPALMNT 1200
Qy 1201 ERTNIQVTVGSSPSP-----LAVVQPAVNIISANGFGFALCOLNV 1243
Db 1201 ERTNIQVTVGSSPSP-----LAVVQPAVNIISANGFGFALCOLNV 1243
Qy 1244 YNVKASSSSRRRSRRIQNOEAFDLVAVKENDDLNHDVLCVTSFGSGRSGMALMEVNL 1303
Db 1244 YNVKASSSSRRRSRRIQNOEAFDLVAVKENDDLNHDVLCVTSFGSGRSGMALMEVNL 1303
Qy 1261 YNVKASSSSRRRSRRIQNOEAFDLVAVKENDDLNHDVLCVTSFGSGRSGMALMEVNL 1320
Db 1261 YNVKASSSSRRRSRRIQNOEAFDLVAVKENDDLNHDVLCVTSFGSGRSGMALMEVNL 1320
Qy 1304 LSGFMVSEALISSEYVAKVYDHGKLNLYDSVNETQFCNIAVNRFKVSNTOADSVS 1363
Db 1304 LSGFMVSEALISSEYVAKVYDHGKLNLYDSVNETQFCNIAVNRFKVSNTOADSVS 1363
Qy 1364 IVDYEEPRQAVRSYNSRVKSSCDLCSVOGCRPCEDGASGSHHSVIFIFCFKLLYF 1423
Db 1364 IVDYEEPRQAVRSYNSRVKSSCDLCSVOGCRPCEDGASGSHHSVIFIFCFKLLYF 1423
Qy 1424 MELNL 1428
Db 1424 MELNL 1428
Qy 1441 MELNL 1445
Db 1441 MELNL 1445

```

RESULT 7  
ABB82168

```

ID ABB82168 standard; Protein; 1445 AA.
XX
AC ABB82168;
XX
DT 23-DEC-2002 (first entry)
XX
DE Human CD109 K1-H7 variant protein.
XX
KW CD109; alpha2 macroglobulin; thioester; cerebroprotective; cardiac;
KW immunosuppressive; haemostatic; anticoagulant; thrombolytic; human;
KW cardiovascular; vasotropic; gene therapy; CD109 K1-H7; variant.
XX
OS Homo sapiens.
XX
PN WO200270696-A2.
XX
PD 12-SEP-2002.
XX
PF 07-MAR-2002; 2002WO-CO00292.
XX
PR 07-MAR-2001; 2001US-273814P.
XX
PA (SCHU/) SCHUH A.
PA (SOTH/) SUTHERLAND R D.
XX
PI Schuh A, Sutherland RD;
XX
DR WPI: 2002-713450/77.
DR N-PSDB; ABQ79967.
XX
PT New CD109 nucleic acids and polypeptides, useful in gene therapy,
PT particularly for treating strokes, myocardial infarctions, thrombosis,
PT thrombocytopenia, autoimmune diseases, or organ or bone marrow
PT transplantation -
XX
PS Claim 21; Fig 3b; 156pp; English.
XX
SS
XX
CC The invention relates to isolated nucleic acid molecules encoding CD109
CC polypeptides. These nucleic acid molecules include the human cDNA
CC sequences comprising CD109 K1, CD109 K1-H7, CD109 K15 or their variants.
CC CD109 is a novel member of the alpha2 macroglobulin (alpha2M)/C3, C4, C5
CC family of thioester-containing proteins. The CD109 polypeptides can be
CC expressed by standard recombinant methodology. The CD109 nucleic acid,
CC CD109 polypeptide or its mimetic, or the anti-CD109 antibody is useful
CC for treating or detecting a disease or disorder, e.g. conditions
CC associated with endothelial activation, platelet activation, activation
CC of the coagulation or fibrinolytic systems, activation of T lymphocytes
CC and of the complement system, quantitative or qualitative abnormalities
CC of platelet function, increased or impaired platelet aggregation and
CC activation, increased or impaired activation of the coagulation and/or
CC fibrinolytic systems, or impaired or increased immune activation. These
CC are also useful for treating cardiovascular disorders, stroke, myocardial
CC infarction, thrombosis, embolism, peripheral vascular disease,
CC thrombocytopenia, thrombocytopenia, autoimmune diseases, organ
CC transplantation, or bone marrow transplantation. The CD109 nucleic acids
CC or their antisense nucleotide sequence are useful in gene therapy for
CC treating these conditions. The present sequence represents the human
CC CD109 K1-H7 variant sequence.
XX
SQ Sequence 1445 AA;
XX
Query Match 99.6%; Score 7317.5; DB 23; Length 1445;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 1426; Conservative 1; Mismatches 1; Indels 17; Gaps 1;
Qy 1 MCGPPLTLTAHLVCCTAALAVAPGPRFLVTAPGIRPGNVITIGVLEHCHPSQVTVKA 60
Db 1 MCGPPLTLTAHLVCCTAALAVAPGPRFLVTAPGIRPGNVITIGVLEHCHPSQVTVKA 60
Qy 61 ELKTSANLTVSVLEAGVVEKSGSKTLTLPSPINSADITIELRYTGRQDELFPSNST 120
Db 61 ELKTSANLTVSVLEAGVVEKSGSKTLTLPSPINSADITIELRYTGRQDELFPSNST 120

```

```

QY 121 RLSFTKRI SVITQTDKALYKPKQVKRIRIVTLFSDPKYKTSNLILIKDKPKSNLIQOML 180
DB 121 RLSFTKRI SVITQTDKALYKPKQVKRIRIVTLFSDPKYKTSNLILIKDKPKSNLIQOML 180
QY 181 SQQSLGVISTKFOGLSHPIIGDMSIQOVVNDQTYQOSQVSEYVLPRFEVTLQPLKCS 240
DB 181 SQQSLGVISTKFOGLSHPIIGDMSIQOVVNDQTYQOSQVSEYVLPRFEVTLQPLKCS 240
QY 241 MNSKHLNGITITAKTYTGKPKVKGDTLTLPLSPFKGKKKNIITKTFKINGSANFENDEEMK 300
DB 241 MNSKHLNGITITAKTYTGKPKVKGDTLTLPLSPFKGKKKNIITKTFKINGSANFENDEEMK 300
QY 301 NMVDSNGISEYLDLSSPGPVEILTTVTESVTGISRNSTNVPEKQDHYIIEFDYTVL 360
DB 301 NMVDSNGISEYLDLSSPGPVEILTTVTESVTGISRNSTNVPEKQDHYIIEFDYTVL 360
QY 361 KPSLNFATVKTADGNOLTLBERRNNVITVTQNTTEWVSGNSNGQKKEAVQKINY 420
DB 361 KPSLNFATVKTADGNOLTLBERRNNVITVTQNTTEWVSGNSNGQKKEAVQKINY 420
QY 421 TVPQSGTFKIEPPILEDSELOLKAYFLGSKSMVAHSLFKSPSKTYIQLKTRDENIKVG 480
DB 421 TVPQSGTFKIEPPILEDSELOLKAYFLGSKSMVAHSLFKSPSKTYIQLKTRDENIKVG 480
QY 481 SPFELVSGNKRKLKELSYMVVRGQLVAVGKONSTMFSLTPEBSMTPRACVIVYYIEDDG 540
DB 481 SPFELVSGNKRKLKELSYMVVRGQLVAVGKONSTMFSLTPEBSMTPRACVIVYYIEDDG 540
QY 541 EIIISVLKIPVOLVKNKIKLWYSKYKAPSEKVSILRISVTPDPSIVGIVAVDKSVNLMN 600
DB 541 EIIISVLKIPVOLVKNKIKLWYSKYKAPSEKVSILRISVTPDPSIVGIVAVDKSVNLMN 600
QY 601 ASNDITMENVVELELYNTGYLLGMPMNSFAVFOCGMLVDNALTKDYIDGVYDNMAY 660
DB 601 ASNDITMENVVELELYNTGYLLGMPMNSFAVFOCGMLVDNALTKDYIDGVYDNMAY 660
QY 661 AERFMEENEGHIVDIDHPSLGSSPHVRKHPETWIMLDTNMGYRIYQREEVTVPSITSW 720
DB 661 AERFMEENEGHIVDIDHPSLGSSPHVRKHPETWIMLDTNMGYRIYQREEVTVPSITSW 720
QY 721 VATGVVISEDGLGLTTTPVELQAFOPFFIPLNLPYSVIRGEEFALETIENYLKDATEV 780
DB 721 VATGVVISEDGLGLTTTPVELQAFOPFFIPLNLPYSVIRGEEFALETIENYLKDATEV 780
QY 781 KVIIEKSDKEDILMTSSEINATGHOOOTLLVSEDEATVLPFRPHLGEIPITWTALSP 840
DB 781 KVIIEKSDKEDILMTSSEINATGHOOOTLLVSEDEATVLPFRPHLGEIPITWTALSP 840
QY 841 ASDAVTOMILVKAEGIEKYSOSIILLDTNRLQSTLKTLSFSPPNVTGSEKRVQITAI 900
DB 841 ASDAVTOMILVKAEGIEKYSOSIILLDTNRLQSTLKTLSFSPPNVTGSEKRVQITAI 900
QY 901 GVDLGPISNGLASLIRMPYGCGBQNMNINPANIYILDYITKKKQUTLNLKEXALSFMKQ 960
DB 901 GVDLGPISNGLASLIRMPYGCGBQNMNINPANIYILDYITKKKQUTLNLKEXALSFMKQ 960
QY 961 YORELLYQREOGSFAFGNYDPSGSTWLSAFLRCFLEADPYIDIDQNVLAHTYTWLKG 1020
DB 961 YORELLYQREOGSFAFGNYDPSGSTWLSAFLRCFLEADPYIDIDQNVLAHTYTWLKG 1020
QY 1021 QKSNGEFMDPGRVHISELQGNKSPVTLTAIVTSLGVRKQPINIDVQESIHFLSEFS 1080
DB 1021 QKSNGEFMDPGRVHISELQGNKSPVTLTAIVTSLGVRKQPINIDVQESIHFLSEFS 1080
QY 1081 RGISNNTYLTALITVALSSVSGPKAKALNMLTWRAEBOEGQMFVWSSEKLSDSWQPSL 1140
DB 1081 RGISNNTYLTALITVALSSVSGPKAKALNMLTWRAEBOEGQMFVWSSEKLSDSWQPSL 1140
QY 1141 DIEVAAYALLSHFLQFQSEGIPIRMWLSRQNSLGGFASPTODTTVALKALSEFALNMT 1200
DB 1141 DIEVAAYALLSHFLQFQSEGIPIRMWLSRQNSLGGFASPTODTTVALKALSEFALNMT 1200
QY 1201 ERTNIQVTVTGPSSPSP-----LAVVQPMVNIISANGFPAICQLNV 1243

```

```

DB 1201 ERTNIQVTVTGPSSPSPFKFLIDTHNRLLOTAELAVVQPMVNIISANGFPAICQLNV 1260
QY 1244 YNVKASGSSRRRRRIQNOEAFDLVAVKENDDIINHVDLWVCSFSGPSGSMALMEVNL 1303
DB 1261 YNVKASGSSRRRRRIQNOEAFDLVAVKENDDIINHVDLWVCSFSGPSGSMALMEVNL 1320
QY 1304 LSGFMVPSSEALISLETYKVEYDHGKLNLYLDSVNETQFCVNIIPAVNFKVSNTOASVS 1363
DB 1321 LSGFMVPSSEALISLETYKVEYDHGKLNLYLDSVNETQFCVNIIPAVNFKVSNTOASVS 1380
QY 1364 IVDYIEPRRQAVRSYNSBVKLSSCDLCSDVQGCPCEDGASGSHHSSVIFIFCKLIYF 1423
DB 1381 IVDYIEPRRQAVRSYNSBVKLSSCDLCSDVQGCPCEDGASGSHHSSVIFIFCKLIYF 1440
QY 1424 MELWL 1428
DB 1441 MELWL 1445

RESULT 8
AA019373
ID AA019373 standard; Protein; 1445 AA.
XX
AC AA019373;
XX
DT 04-DEC-2002 (first entry)
XX
DE Human platelet alloantigen Govb.
XX
KW Human; platelet alloantigen; Gova; Govb; single nucleotide polymorphism;
KW SNP; biallelic; bleeding disorder; post-transfusion purpura;
KW post-transfusion platelet refractoriness; haemostatic; vaccine;
KW neonatal alloimmune thrombocytopenia.
XX
OS Homo sapiens.
XX
PN MO20020738-A2.
XX
PD 12-SEP-2002.
XX
PF 07-MAR-2002; 2002WO-CA00291.
XX
PR 07-MAR-2001; 2001US-273941P.
XX
PA (SCHU/) SCHUH A.
XX
PI Schuh A, Ouwehand W;
XX
DR MPI; 2002-713460/77.
XX
DR N-PSDB; AAL49816.
XX
PT New isolated oligonucleotide binding to a region of CD109 nucleic acid
PT having a single nucleotide polymorphism that distinguishes a Gova
PT and/or Govb allele, useful for treating blood disorders e.g. alloimmune
PT thrombocytopenia
XX
PS Disclosure; Page 41-47; 69pp; English.
XX
CC The present invention relates to a sequence capable of binding
CC specifically to a CD109 nucleic acid which has a single nucleotide
CC polymorphism that distinguishes the Gova and Govb alleles. Detection of
CC the Gov genotype is useful for detecting whether the subject has or is at
CC risk of a blood disease, disorder or abnormal physical state, such as
CC bleeding, or increased risk of bleeding, due to alloimmune destruction of
CC blood platelets, e.g., post-transfusion purpura, post-transfusion
CC platelet refractoriness or neonatal alloimmune thrombocytopenia. The
CC nucleic acid and polypeptide are useful for Gov genotyping or phenotyping
CC individuals. The present sequence is the Govb encoded protein.
XX
SQ Sequence 1445 AA;

```

Query Match 99.6%; Score 7317.5; DB 23; Length 1445;

Best Local Similarity 98.7%; Pred. No. 0;  
Matches 1426; Conservative 1; Mismatches 1; Indels 17; Gaps 1;

```

QY 1 MGGPPLTAAHLTCCTAALAVAPGPPLVAPGIRPGNVITGVLELHCPQVYVKA 60
DB 1 MGGPPLTAAHLTCCTAALAVAPGPPLVAPGIRPGNVITGVLELHCPQVYVKA 60
QY 61 ELKTAANLTVSVLEAGVPEKSPKTLTLPSPLSNADELYELVTKGTODEILFNSNT 120
DB 61 ELKTAANLTVSVLEAGVPEKSPKTLTLPSPLSNADELYELVTKGTODEILFNSNT 120
QY 121 RLSEETKRISVFICTDKALYKPKQVKRYITLPSDPKPYKTSNILLIKPKSNLLOQWL 180
DB 121 RLSEETKRISVFICTDKALYKPKQVKRYITLPSDPKPYKTSNILLIKPKSNLLOQWL 180
QY 121 RLSEETKRISVFICTDKALYKPKQVKRYITLPSDPKPYKTSNILLIKPKSNLLOQWL 180
DB 121 RLSEETKRISVFICTDKALYKPKQVKRYITLPSDPKPYKTSNILLIKPKSNLLOQWL 180
QY 181 SQGSDLVISKTFOLSHPILGDMISIQVQVNDQTYYSQVSEVYLKPEFVTLQTPLYCS 240
DB 181 SQGSDLVISKTFOLSHPILGDMISIQVQVNDQTYYSQVSEVYLKPEFVTLQTPLYCS 240
QY 241 MNSHGLNGITTAKYTYGKPKVADVTLPPLSPFWGKKKNTKTEPKINGSANFSNDEEMK 300
DB 241 MNSHGLNGITTAKYTYGKPKVADVTLPPLSPFWGKKKNTKTEPKINGSANFSNDEEMK 300
QY 301 NVMSSSNGLSSEYLDLSSPGPVEILTTYTESVTGISRNVSTNVPKQHDYIIIEFDYTTVL 360
DB 301 NVMSSSNGLSSEYLDLSSPGPVEILTTYTESVTGISRNVSTNVPKQHDYIIIEFDYTTVL 360
QY 361 KPSLNFTATVAVTRADGNQLTLERRNNVITVQRYATEWGSNGSNOMKQVAKINY 420
DB 361 KPSLNFTATVAVTRADGNQLTLERRNNVITVQRYATEWGSNGSNOMKQVAKINY 420
QY 421 TVPQSGTFKIEPPILEDSESLQAKAYFLGSKSNAVHSLFKSPSKTYIQLKTRDENIKVG 480
DB 421 TVPQSGTFKIEPPILEDSESLQAKAYFLGSKSNAVHSLFKSPSKTYIQLKTRDENIKVG 480
QY 481 SPPELVVSGNRLKELSYMVSRCQLYAVGKONSTMSLTPENSWTPKACIVYIYIEDDG 540
DB 481 SPPELVVSGNRLKELSYMVSRCQLYAVGKONSTMSLTPENSWTPKACIVYIYIEDDG 540
QY 541 EIIISDVLIKIPQVAFKPKIKLYMSKVAEPEKSLHSVQPSIYGIYAVDVSVLNM 600
DB 541 EIIISDVLIKIPQVAFKPKIKLYMSKVAEPEKSLHSVQPSIYGIYAVDVSVLNM 600
QY 601 ASNDITMENVHELELYNTGYLLGMFMSFAVFOECGLMVLTDANLTKDYIDGYDAEY 660
DB 601 ASNDITMENVHELELYNTGYLLGMFMSFAVFOECGLMVLTDANLTKDYIDGYDAEY 660
QY 661 AERMEENEGHIVDIHDSLGSSPHVRKHPEPTWIMLDTNMGRYIYQEFVTVPDSITSW 720
DB 661 AERMEENEGHIVDIHDSLGSSPHVRKHPEPTWIMLDTNMGRYIYQEFVTVPDSITSW 720
QY 721 VATGFVISEDGLGLTTTPVELQAFOPPEFLNPSYVINGEERFALITTFNYLKDADEV 780
DB 721 VATGFVISEDGLGLTTTPVELQAFOPPEFLNPSYVINGEERFALITTFNYLKDADEV 780
QY 781 KVIIEKSDKPDILMTSSEINATGHQQLTVSESDATVLPPIRTHGEIPIYVTALESPT 840
DB 781 KVIIEKSDKPDILMTSSEINATGHQQLTVSESDATVLPPIRTHGEIPIYVTALESPT 840
QY 841 ASDAVTQMLIVKAGIEKYSQSILLDTNRLQSTLKTLSFSPPTVTVGSEKVOITAI 900
DB 841 ASDAVTQMLIVKAGIEKYSQSILLDTNRLQSTLKTLSFSPPTVTVGSEKVOITAI 900
QY 901 GDVIGPSINGLASLIRMPYGGGEONMINFAPNIYILDYLTCKQLTNLEKALSPKROG 960
DB 901 GDVIGPSINGLASLIRMPYGGGEONMINFAPNIYILDYLTCKQLTNLEKALSPKROG 960
QY 961 YQRELLAQREDSGSAFGNDPSGSTMASAVLACFLEADPYIDIDQNVLHRTYTWLKGH 1020
DB 961 YQRELLAQREDSGSAFGNDPSGSTMASAVLACFLEADPYIDIDQNVLHRTYTWLKGH 1020
QY 1021 QKSNGEFMDPGRVHISELOGGNKSPVTLTAVYVTSLLGKRYQPNIDVQSEIHLESEFS 1080
DB 1021 QKSNGEFMDPGRVHISELOGGNKSPVTLTAVYVTSLLGKRYQPNIDVQSEIHLESEFS 1080

```

```

DB 1021 QKSNGEFMDPGRVHISELOGGNKSPVTLTAVYVTSLLGKRYQPNIDVQSEIHLESEFS 1080
QY 1081 RGISDNVYTLALITYALSSVGSFPKAEALNMLTWRAEOEGMGQFWVSESKLSDSQPRSL 1140
DB 1081 RGISDNVYTLALITYALSSVGSFPKAEALNMLTWRAEOEGMGQFWVSESKLSDSQPRSL 1140
QY 1141 DIFVAAVYALISHFLQOTSSEGIPIMKWSRORNSLGGFASSTODTYVALKALSEFALMNT 1200
DB 1141 DIFVAAVYALISHFLQOTSSEGIPIMKWSRORNSLGGFASSTODTYVALKALSEFALMNT 1200
QY 1201 ERTNIQVTVGSPSPSP-----LAVQPMVAVNISANGGFALICOLNVY 1243
DB 1201 ERTNIQVTVGSPSPSPVKFLDTHNRLLQTNELAVQPMVAVNISANGGFALICOLNVY 1260
QY 1244 YNFKASGSSRRRSRISQNOEAFDLDAVAKENKDLNHDVNDVCTSFSGPGRSGMALMEVNL 1303
DB 1261 YNFKASGSSRRRSRISQNOEAFDLDAVAKENKDLNHDVNDVCTSFSGPGRSGMALMEVNL 1320
QY 1304 LSGFMPVSEALISLSEFTVKYEDHGLNLYLDSVNETQCVNIPAVRNKVSNTQDASVS 1363
DB 1321 LSGFMPVSEALISLSEFTVKYEDHGLNLYLDSVNETQCVNIPAVRNKVSNTQDASVS 1380
QY 1364 IVDYEPBRQAVRSVSEVKLSGCDLCSDVQGRPCEDGASGSHHSVYIFCFPLLYF 1423
DB 1381 IVDYEPBRQAVRSVSEVKLSGCDLCSDVQGRPCEDGASGSHHSVYIFCFPLLYF 1440
QY 1424 MELML 1428
DB 1441 MELML 1445

RESULT 9
AAB12127
ID AAB12127 standard; Protein; 1445 AA.
XX
AC AAB12127;
XX
DT 02-FEB-2001 (first entry)
XX
DE Hydrophobic domain protein isolated from HT-1080 cells.
XX
KW Human; secreted protein; membrane protein; hydrophobic domain;
KW proliferation control; differentiation induction; material transport;
KW biophysical; signal receptor; ion channel; transporter; immunostimulant;
KW immunosuppressant; haematopoiesis regulator; chemotactic; chemokinetic;
KW haemostatic; thrombolytic; anti-inflammatory; tumour inhibition;
KW autoimmune disease; Alzheimer's disease; Parkinson's disease; cancer.
XX
OS Homo sapiens.
XX
PN MO200029448-A2.
XX
PD 25-MAY-2000.
XX
PE 17-NOV-1999; 99MO-JP06412.
XX
PR 17-NOV-1998; 98JP-0326255.
PR 22-DEC-1998; 98JP-0364315.
PR 16-MAR-1999; 99JP-0069811.
PR 27-APR-1999; 99JP-0119299.
PR 19-MAY-1999; 99JP-0138169.
XX
PA (SAGA) SAGAMI CHEM RES CENT.
PA (PROT-) PROTEGENE INC.
XX
PI Kato S, Kimura T;
XX
DR N-PSDB; AAA60199, AAA62010.
DR
PT Proteins comprising hydrophobic regions, such as secretory and membrane
PT proteins, useful in research and diagnostics and having various
PT activities e.g. immunomodulatory, antiinflammatory, chemokinetic,

```



PN WO200285942-A2.  
XX 31-OCT-2002.  
XX 24-APR-2002; 2002WO-CA00560.  
XX 24-APR-2001; 2001US-285713P.  
XX 14-FEB-2002; 2002US-356163P.  
XX (UMC-) UNIV MCGILL.  
XX Philip A, Tam B;  
XX WPI; 2003-093100/08.  
XX N-PSDB; AAD49436.  
XX  
XX Novel transforming growth factor (TGF)-beta 1 binding reagent which  
XX comprises r150 protein which acts as accessory receptor of TGF-beta,  
XX useful for negatively modulating TGF-beta activity, and thus for  
XX treating cancer -  
XX  
XX Claim 1; Page 112-117; 127pp; English.  
XX  
XX The invention relates to novel transforming growth factor (TGF)-beta1  
XX binding reagent which comprises a glycosylphosphatidy1 inositol (GPI)-  
XX anchored TGF-beta1 binding protein referred to as r150 which acts as  
XX an accessory receptor of TGF-beta. The invention is used for negatively  
XX modulating TGF-beta activity, and thus for treating conditions  
XX characterised by overproduction of TGF-beta, such as cancer. Antisense  
XX molecules of the invention are used for increasing TGF-beta availability  
XX and increase graft success. The present sequence is human blood cell  
XX surface antigen, CD109. CD109 is a variant of r150 sequence.  
SQ Sequence 1382 AA;

Query Match 96.4%; Score 7081; DB 24; Length 1382;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1379; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 22 VAPGPRFLVTAGRIIRPGNVITIGVLELHPCSQVTVKAEILKTASNLTVSLAEAGVFE 81  
DB 1 VAPGPRFLVTAGRIIRPGNVITIGVLELHPCSQVTVKAEILKTASNLTVSLAEAGVFE 60  
QY 82 KGSFKTLLPSLPLNSADEIYELVGTODEILFNSSTRLSFETKRISVFIQTDKALYK 141  
DB 61 KGSFKTLLPSLPLNSADEIYELVGTODEILFNSSTRLSFETKRISVFIQTDKALYK 120  
QY 142 PKQEVKFRIVTLFSDPKRYKTSNLILIKDPKSNLIQOMLSQOSDLGVISKTFOUSSHPIL 201  
DB 121 PKQEVKFRIVTLFSDPKRYKTSNLILIKDPKSNLIQOMLSQOSDLGVISKTFOUSSHPIL 180  
QY 202 GDMGIVOVNDQTYQSFQVSEYVLPKFEVTLQTPLCVSNMSKILNCTITAKYVYGPVK 261  
DB 181 GDMGIVOVNDQTYQSFQVSEYVLPKFEVTLQTPLCVSNMSKILNCTITAKYVYGPVK 240  
QY 262 GGVLTFLPLSPFKGKKKITFTFKINGSANFSFNDDEKKNVMSNGISEYLDLSSGPV 321  
DB 241 GGVLTFLPLSPFKGKKKITFTFKINGSANFSFNDDEKKNVMSNGISEYLDLSSGPV 300  
QY 322 EILTTVTESVTGISRNSTNVFQKHDIYIBFDYTVLAKPSLNFATVKTTRADGNQLT 381  
DB 301 EILTTVTESVTGISRNSTNVFQKHDIYIBFDYTVLAKPSLNFATVKTTRADGNQLT 360  
QY 382 LEEERNNVVITVQNTTEYWSGNSGNQKMEAVQKINITYPOSGTFKRIEPILEDSESL 441  
DB 361 LEEERNNVVITVQNTTEYWSGNSGNQKMEAVQKINITYPOSGTFKRIEPILEDSESL 420  
QY 442 QLKAYFLGSKSSMAVHSLFKSPSKTYIQKTRDNKIKVGSFPELVGSKRLKELSTYVV 501  
DB 421 QLKAYFLGSKSSMAVHSLFKSPSKTYIQKTRDNKIKVGSFPELVGSKRLKELSTYVV 480  
QY 502 SRGQLVAVGKONSTMFSLTPENSWTPKACVIVYVIEDDEGIISDVLKIPQVLFKNKIKL 561  
DB 502 SRGQLVAVGKONSTMFSLTPENSWTPKACVIVYVIEDDEGIISDVLKIPQVLFKNKIKL 561

DB 481 SRGQLVAVGKONSTMFSLTPENSWTPKACVIVYVIEDDEGIISDVLKIPQVLFKNKIKL 540  
QY 562 YMSKVAAEPESEKSLRISVTPQDSIVGIVAVDKSVLMMASNDITMENNVHELELYNTGY 621  
DB 541 YMSKVAAEPESEKSLRISVTPQDSIVGIVAVDKSVLMMASNDITMENNVHELELYNTGY 600  
QY 622 YLGMFMSFAVFOECGLMWLTDLNLTQDYIDGYDNAEYAEERMEENEGHIVDIHDFSLG 681  
DB 601 YLGMFMSFAVFOECGLMWLTDLNLTQDYIDGYDNAEYAEERMEENEGHIVDIHDFSLG 660  
QY 682 SSPHVRKHPEETIMWLTQNGVRIYQEFETVPDSITSWATGFGVISEDGLGTLTPVE 741  
DB 661 SSPHVRKHPEETIMWLTQNGVRIYQEFETVPDSITSWATGFGVISEDGLGTLTPVE 720  
QY 742 LQAFQPFIFLNPYSVIRGEFEALITTFNVLKDADEVVYIEKSDKPDILMTSSEINA 801  
DB 721 LQAFQPFIFLNPYSVIRGEFEALITTFNVLKDADEVVYIEKSDKPDILMTSSEINA 780  
QY 802 TGHQQTLLVPSDEGATVLPFIRPTHGELPIYVYALSPTASDAVTOMLYKAGIEKSYG 861  
DB 781 TGHQQTLLVPSDEGATVLPFIRPTHGELPIYVYALSPTASDAVTOMLYKAGIEKSYG 840  
QY 862 QSTILLDTNRLQSTLTKTSFSFPPTVTVSGSEEVQITAGDVLGSPSLNGSLIRMPYGC 921  
DB 841 QSTILLDTNRLQSTLTKTSFSFPPTVTVSGSEEVQITAGDVLGSPSLNGSLIRMPYGC 900  
QY 922 GEONMNFAPNIYILDYLTQKKQLTDLNLEKALSPRQGYORELYQREDSFSARGNYD 981  
DB 901 GEONMNFAPNIYILDYLTQKKQLTDLNLEKALSPRQGYORELYQREDSFSARGNYD 960  
QY 982 PGSGTWLSAFVLCFLPADBYIDIDQNVLHRTYTWLKHQKNGEWFDPGRVHSELQGG 1041  
DB 961 PGSGTWLSAFVLCFLPADBYIDIDQNVLHRTYTWLKHQKNGEWFDPGRVHSELQGG 1020  
QY 1042 NKSPVTLTAITYSLGKRYQPNIDVQSIHFLSEFSRGISDNTLALITALSYSYG 1101  
DB 1021 NKSPVTLTAITYSLGKRYQPNIDVQSIHFLSEFSRGISDNTLALITALSYSYG 1080  
QY 1102 PKAKEALNMLTWRAEOEGQMFVWSESKLSDMQRSLDIEVAAYALSHFLQFQTSSEG 1161  
DB 1081 PKAKEALNMLTWRAEOEGQMFVWSESKLSDMQRSLDIEVAAYALSHFLQFQTSSEG 1140  
QY 1162 IPIRMLSRQNSLGGFASDTTVALKALSEFAALMTERINIQTVTGSSSPSLAVV 1221  
DB 1141 IPIRMLSRQNSLGGFASDTTVALKALSEFAALMTERINIQTVTGSSSPSLAVV 1200  
QY 1222 QPMANVISANGFPAICQNLVYVNVKASGSSRRRSIQNOEAPDLDVAVKENDDLNHYD 1281  
DB 1201 QPMANVISANGFPAICQNLVYVNVKASGSSRRRSIQNOEAPDLDVAVKENDDLNHYD 1260  
QY 1282 LNVCTSPSGRSGMALMEVNLISGFMPVSEALISSETVAKVEYDQKTLNLYDSVNETQ 1341  
DB 1261 LNVCTSPSGRSGMALMEVNLISGFMPVSEALISSETVAKVEYDQKTLNLYDSVNETQ 1320  
QY 1342 FCVNIQAVNFKVSNTOQASVSIVDYEBPRQAVRSYNEVKSCLDCSDVQGCRRPED 1401  
DB 1321 FCVNIQAVNFKVSNTOQASVSIVDYEBPRQAVRSYNEVKSCLDCSDVQGCRRPED 1380  
QY 1402 GA 1403  
DB 1381 GA 1382

RESULT 11  
AAE32015 standard; Protein; 1399 AA.  
ID AAE32015;  
XX AAE32015;  
AC AAE32015;  
XX 24-MAR-2003 (first entry)  
DT Human r150 protein #3.  
XX DE

KW Human; GPI-anchored TGF-beta1 binding protein; r150 protein; cancer;  
 KW glycosylphosphatidylinositol; transforming growth factor-beta1;  
 KW therapy.  
 XX  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT Misc-difference 682 /label= Ser, Tyr  
 FT /note= "Encoded by TNC"  
 XX  
 PN MO200285942-A2.  
 XX  
 PD 31-OCT-2002.  
 XX  
 PF 24-APR-2002; 2002MO-CA00560.  
 XX  
 PR 24-APR-2001; 2001US-285713P.  
 PR 14-FEB-2002; 2002US-356163P.  
 XX  
 PA (UYN-) UNIV MCGILL.  
 XX  
 PI Philip A. Tam B;  
 DR WPI; 2003-093100/08.  
 DR N-PSDB; AAD49437.  
 XX  
 PT Novel transforming growth factor (TGF)-beta 1 binding reagent which  
 PT comprises r150 protein which acts as accessory receptor of TGF-beta,  
 PT useful for negatively modulating TGF-beta activity, and thus for  
 PT treating cancer -  
 XX  
 PS Claim 1; Page 120-126; 127pp; English.  
 XX  
 CC The invention relates to novel transforming growth factor (TGF)-beta1  
 CC binding reagent which comprises a glycosylphosphatidylinositol (GPI)-  
 CC anchored TGF-beta1 binding protein referred to as r150 which acts as  
 CC an accessory receptor of TGF-beta. The invention is used for negatively  
 CC modulating TGF-beta activity, and thus for treating conditions  
 CC characterised by overproduction of TGF-beta, such as cancer. Antisense  
 CC molecules of the invention are used for increasing TGF-beta availability  
 CC and increase graft success. The present sequence is human r150 protein.  
 CC  
 SQ Sequence 1399 AA;  
 Query Match 96.2%; Score 7068.5; DB 24; Length 1399;  
 Best Local Similarity 98.6%; Pred. No. 0;  
 Matches 1380; Conservative 1; Mismatches 1; Indels 17; Gaps 1;

QY 382 IEERRNNVITVTOHNTYEWSGNSGNQKMEAVKINYTVPOSGTFKIEPIEDSSSEL 441  
 DB 361 IEERRNNVITVTOHNTYEWSGNSGNQKMEAVKINYTVPOSGTFKIEPIEDSSSEL 420  
 QY 442 QLKAYFLGSKSMVAHSLPKSPSKTYIQLKTRDENIKVGSPEELVSGNKKLKEISYMW 501  
 DB 421 QLKAYFLGSKSMVAHSLPKSPSKTYIQLKTRDENIKVGSPEELVSGNKKLKEISYMW 480  
 QY 502 SRGOLVAAGKONSTWPSLTPEMSWTPKACVITYIYEDGEIISDLKIPVOLFKNKIKL 561  
 DB 481 SRGOLVAAGKONSTWPSLTPEMSWTPKACVITYIYEDGEIISDLKIPVOLFKNKIKL 540  
 QY 562 YMSKYKABPSEKVSIRISVTOPODSIVGIVAVDKSVNLMAASNDITMENVHELELYNTGY 621  
 DB 541 YMSKYKABPSEKVSIRISVTOPODSIVGIVAVDKSVNLMAASNDITMENVHELELYNTGY 600  
 QY 601 YLGMFMNSFAVFOEGCLMWLTDANTKDYIDGVYDNAMEYARFMEENEGHIVDIDHPSLG 660  
 DB 682 SEPHVKHPPETWIMLDITNMGRIVQEFVTPVPSITSWVATNGFVISEDIGLITTPVE 741  
 DB 661 SEPHVKHPPETWIMLDITNMGRIVQEFVTPVPSITSWVATNGFVISEDIGLITTPVE 720  
 QY 742 LQAFQPFIFLNLPSYVIRGEFALEITIFNYLKDATERVKYIIEKSDKFDILMTSEINA 801  
 DB 721 LQAFQPFIFLNLPSYVIRGEFALEITIFNYLKDATERVKYIIEKSDKFDILMTSEINA 780  
 QY 802 TGHQOTLLVPSDEGATVLPPIRPHLGEIPITVTALSPASDAVTOMLVYAEIGIEKYS 861  
 DB 781 TGHQOTLLVPSDEGATVLPPIRPHLGEIPITVTALSPASDAVTOMLVYAEIGIEKYS 840  
 QY 862 OSILDLTDNRLOSTLKLSPSPNNYTGSEBVOITLIGVLAGSINGLASITMPGC 921  
 DB 841 OSILDLTDNRLOSTLKLSPSPNNYTGSEBVOITLIGVLAGSINGLASITMPGC 900  
 QY 922 GEONMINFAPNIYIDVLTKKKQOLDNLKERALSMPRGYORELLYQREDSFSAFGYND 981  
 DB 901 GEONMINFAPNIYIDVLTKKKQOLDNLKERALSMPRGYORELLYQREDSFSAFGYND 960  
 QY 962 PGSGTWLSAFVLRCELEADPYIIDONVLAHRTYTWLKGHOKSNEGFMDPGRVHSELQGG 1041  
 DB 961 PGSGTWLSAFVLRCELEADPYIIDONVLAHRTYTWLKGHOKSNEGFMDPGRVHSELQGG 1020  
 QY 1042 NKSPVTLTAIYVTSILGTRKYQPNIDVQESIHFLSEFSRGISDNYTALITTAISVGS 1101  
 DB 1021 NKSPVTLTAIYVTSILGTRKYQPNIDVQESIHFLSEFSRGISDNYTALITTAISVGS 1080  
 QY 1102 PKAKALNMLTWRAEOEGMOPWSSSEKLSDSMOPRLDIEVAVALLSHFLOPOTSEG 1161  
 DB 1081 PKAKALNMLTWRAEOEGMOPWSSSEKLSDSMOPRLDIEVAVALLSHFLOPOTSEG 1140  
 QY 1162 IPIRMWLSRORNSLGGFASTODTTVALKALSEFALMTERNTNIQVTVGPSSBSP---- 1217  
 DB 1141 IPIRMWLSRORNSLGGFASTODTTVALKALSEFALMTERNTNIQVTVGPSSBSP---- 1200  
 QY 1218 -----LAVOPMAVINISANGFGAICQANVYVNVKASGSSRRRRRSTIONDAF 1264  
 DB 1201 IDTHNRLLAQTAELVAVOPMAVINISANGFGAICQANVYVNVKASGSSRRRRRSTIONDAF 1260  
 QY 1265 DLDVAVKENKODLNVNDLVNCTSSFGPGRSGMALMEVVLGSPFWPSAISLSEFVKVYE 1324  
 DB 1261 DLDVAVKENKODLNVNDLVNCTSSFGPGRSGMALMEVVLGSPFWPSAISLSEFVKVYE 1320  
 QY 1325 YDHGRLNLYDSVNETQCVNIIPAVRNKVSNTODASVIVDYEPRROAVRSYNSBYKL 1384  
 DB 1321 YDHGRLNLYDSVNETQCVNIIPAVRNKVSNTODASVIVDYEPRROAVRSYNSBYKL 1380  
 QY 1385 SSCDLCSVQGRCPEDGA 1403  
 DB 1381 SSCDLCSVQGRCPEDGA 1399



RESULT 12  
ID ABB82169 standard; Protein; 1067 AA.  
AC ABB82169;  
XX  
XX  
XX 23-DEC-2002 (first entry)  
DE Human CD109 K15 protein.  
XX  
XX CD109; alpha2 macroglobulin; thioester; cerebroprotective; cardiac; immunosuppressive; haemostatic; anticoagulant; thrombolytic; human; cardiovascular; vasotropic; gene therapy; CD109 K15.  
XX  
XX Homo sapiens.  
XX  
XX MO200270696-A2.  
XX  
XX 12-SEP-2002.  
XX  
XX 07-MAR-2002; 2002MO-CA00292.  
XX  
XX 07-MAR-2001; 2001US-273814P.  
XX  
XX (SCHU/) SCHUH A.  
XX (SUTH/) SUTHERLAND R D.  
XX  
XX Schuh A, Sutherland RD;  
XX  
XX WPI; 2002-713450/77.  
XX  
XX N-PSDB; ABQ79968.  
XX  
XX New CD109 nucleic acids and polypeptides, useful in gene therapy, particularly for treating strokes, myocardial infarctions, thrombosis, thrombocytopenia, autoimmune diseases, or organ or bone marrow transplantation -  
XX  
XX  
XX Claim 21; Fig 5a; 156p; English.  
XX  
XX The invention relates to isolated nucleic acid molecules encoding CD109 polypeptides. These nucleic acid molecules include the human cDNA sequences comprising CD109 K1, CD109 K1-H7, CD109 K15 or their variants. CD109 is a novel member of the alpha2 macroglobulin (alpha2M)/C3, C4, C5 family of thioester-containing proteins. The CD109 polypeptides can be expressed by standard recombinant methodology. The CD109 nucleic acid, CD109 polypeptide or its mimetic, or the anti-CD109 antibody is useful for treating or detecting a disease or disorder, e.g. conditions associated with endothelial activation, platelet activation, activation of the coagulation or fibrinolytic systems, quantitative or qualitative abnormalities of platelet function, increased or impaired platelet aggregation and activation, increased or impaired activation of the coagulation and/or fibrinolytic systems, or impaired or increased immune activation. These are also useful for treating cardiovascular disorders, stroke, myocardial infarction, thrombosis, embolism, peripheral vascular disease, thrombocytopenia, thrombocytopenia, autoimmune diseases, organ transplantation, or bone marrow transplantation. The CD109 nucleic acids or their antisense nucleotide sequence are useful in gene therapy for treating these conditions. The present sequence represents the human CD109 K15 sequence.  
XX  
XX  
SQ Sequence 1067 AA;

Query Match 74.4%; Score 5470; DB 23; Length 1067;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1062; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGGPPLTAAHLTCCTAALAVAGPRFLVTAAGIRPGNVITGVLELHCPQVTVKA 60  
DB 1 MGGPPLTAAHLTCCTAALAVAGPRFLVTAAGIRPGNVITGVLELHCPQVTVKA 60  
QY 61 ELKTASNLTVSVLEAGVFEKGSFKTLTLPSPILNSADEIYELRVGTODEILFNSNT 120

DB 61 ELKTASNLTVSVLEAGVFEKGSFKTLTLPSPILNSADEIYELRVGTODEILFNSNT 120  
QY 121 RLSFETKRISVFIOTDKALYKPKQEVKFRIVTLTSPDPKPKTSLNLIKDPKSNLIQOML 180  
DB 121 RLSFETKRISVFIOTDKALYKPKQEVKFRIVTLTSPDPKPKTSLNLIKDPKSNLIQOML 180  
QY 181 SQQSDGVISKTFOSSHPIIGDWSIQOVVNDQTYQSFQVSEYVLPKFEVTLQTPLYCS 240  
DB 181 SQQSDGVISKTFOSSHPIIGDWSIQOVVNDQTYQSFQVSEYVLPKFEVTLQTPLYCS 240  
QY 241 MNSKHNGTITTAQYTGKPRKGDVTLTPLPLSPWGKKXNTTKPKFNGSNFSPNDEMK 300  
DB 241 MNSKHNGTITTAQYTGKPRKGDVTLTPLPLSPWGKKXNTTKPKFNGSNFSPNDEMK 300  
QY 301 NWDSSNGLSEYDLSSPGVEILTTVTESVITGISHNVSTNVFPKQHDYIIEFFDYTVL 360  
DB 301 NWDSSNGLSEYDLSSPGVEILTTVTESVITGISHNVSTNVFPKQHDYIIEFFDYTVL 360  
QY 361 KPSLNFATATVKTADGNQTLTEERRNNVITYTQNNYTEWGSNGNQKMAVOKINY 420  
DB 361 KPSLNFATATVKTADGNQTLTEERRNNVITYTQNNYTEWGSNGNQKMAVOKINY 420  
QY 421 TVPQSGFKIEPILDDSELOKAYFLGSKSMAVHSLFKSPSKTYIOLKTRDENIKVG 480  
DB 421 TVPQSGFKIEPILDDSELOKAYFLGSKSMAVHSLFKSPSKTYIOLKTRDENIKVG 480  
QY 481 SPFELVVGSKRKELESYVWVSQGLVAVGKONSTFWSLTPENSWTPKACVITYYIEDDG 540  
DB 481 SPFELVVGSKRKELESYVWVSQGLVAVGKONSTFWSLTPENSWTPKACVITYYIEDDG 540  
QY 541 EITSDVLKIPVQLVFNKIKLYWSKYKABSEKVSIRISVTPDSIVGIVAVDKSVNLNM 600  
DB 541 EITSDVLKIPVQLVFNKIKLYWSKYKABSEKVSIRISVTPDSIVGIVAVDKSVNLNM 600  
QY 541 EITSDVLKIPVQLVFNKIKLYWSKYKABSEKVSIRISVTPDSIVGIVAVDKSVNLNM 600  
QY 601 ASDNITMENVHELELYNCGYVYLGEMNSPAVQEGWLTDANLTXYQIDGVYNAEY 660  
DB 601 ASDNITMENVHELELYNCGYVYLGEMNSPAVQEGWLTDANLTXYQIDGVYNAEY 660  
QY 661 AERFMEBEGHIVDIDHFSIGSSPHVRKHPETWIMLDTNMGRIYQEPFVTPDSITSW 720  
DB 661 AERFMEBEGHIVDIDHFSIGSSPHVRKHPETWIMLDTNMGRIYQEPFVTPDSITSW 720  
QY 721 VATGFIYSEDGLGLTTTPVELOAFQPFILNLPYSVIRGEEFALBITFNYLKATEV 780  
DB 721 VATGFIYSEDGLGLTTTPVELOAFQPFILNLPYSVIRGEEFALBITFNYLKATEV 780  
QY 781 KVITIEKSDKRDILMTSEINATGHQOTLLVPSDGAIVLPPIRPTLGEIPITVTLSP 840  
DB 781 KVITIEKSDKRDILMTSEINATGHQOTLLVPSDGAIVLPPIRPTLGEIPITVTLSP 840  
QY 841 ASDAVTOMILVAKEGIEKYSQSILDLTDNRLOSTLTKLSFSFPNTVTGSERVQITAI 900  
DB 841 ASDAVTOMILVAKEGIEKYSQSILDLTDNRLOSTLTKLSFSFPNTVTGSERVQITAI 900  
QY 901 GDVLGFSINGLASLIMPYGCGEONNINFAPNIYILDYITKKKQLTDNLKEKALSPMRQ 960  
DB 901 GDVLGFSINGLASLIMPYGCGEONNINFAPNIYILDYITKKKQLTDNLKEKALSPMRQ 960  
QY 961 YQRELLYQREDSFSAFGNYDPGSGTWLSAFVLRCLPADPYIDQNVLHRTYWLKGH 1020  
DB 961 YQRELLYQREDSFSAFGNYDPGSGTWLSAFVLRCLPADPYIDQNVLHRTYWLKGH 1020  
QY 1021 QKNGSEFMDPGRVTHSELQGNKSPVTLTAIYVTSLLGYRKYQ 1063  
DB 1021 QKNGSEFMDPGRVTHSELQGNKSPVTLTAIYVTSLLGYRKYQ 1063

RESULT 13  
ID ABB82170 standard; Protein; 1067 AA.  
XX ABB82170;  
XX  
XX ABB82170;



therapy; blood cell surface antigen, CD109.  
Homo sapiens.  
MO200285942-A2.  
31-OCT-2002.  
24-APR-2002; 2002MO-CA00560.  
24-APR-2001; 2001US-285713P.  
14-FEB-2002; 2002US-356163P.  
(UTMC-) UNIV MCGILL.  
Philip A, Tam B;  
WPI; 2003-093100/08.  
Novel transforming growth factor (TGF)-beta 1 binding reagent which comprises r150 protein which acts as accessory receptor of TGF-beta, useful for negatively modulating TGF-beta activity, and thus for treating cancer -  
disclosure; Fig 17, 127pp; English.

The invention relates to novel transforming growth factor (TGF)-beta1 binding reagent which comprises a glycosylphosphatidylinositol (GPI)-anchored TGF-beta binding protein referred to as r150 which acts as an accessory receptor of TGF-beta. The invention is used for negatively modulating TGF-beta activity, and thus for treating conditions characterized by overproduction of TGF-beta, such as cancer. Antisense molecules of the invention are used for increasing TGF-beta availability and increase graft success. The present sequence is human blood cell surface antigen, CD109. CD109 is a variant of r150 sequence.  
Note: This sequence is stated to be encoded by human CD109 DNA shown in Figure 17 of the specification, but this does not appear to be the case.

SO Sequence 1445 AA;

Query Match 38.9%; Score 2860.5; DB 24; Length 1445;

Best Local Similarity 59.9%; Pred. No. 3.3e-178;

Matches 861; Conservative 1; Mismatches 559; Indels 17; Gaps 1;

QY 1 MGGPPLTAAHLCTCTAALAVAPGPRFLVTAAGIIRGAVNTIGVELLEHCPQVYKA 60  
DB 1 MGGPPLTAAHLCTCTAALAVAPGPRFLVTAAGIIRGAVNTIGVELLEHCPQVYKA 60  
QY 61 ELKLTASNLTVSVLEAGVPEKSGFKTLTLPPLNSADEIYELRVGTODEILFNSNT 120  
DB 61 GLILTLASNLTVSVLEAGVPEKSGFKTLTLPPLNSADEIYELRVGTODEILFNSNT 120  
QY 121 RLSEETKRIISVFIOTDKALYKPOEVKRYITLPSDFPKYTSNLIKDPKSNLIQOML 180  
DB 121 RLSEETKRIISVFIOTDKALYKPOEVKRYITLPSDFPKYTSNLIKDPKSNLIQOML 180  
QY 121 ALSGCTLAISVPDGTALALTLPLGGLVPAITLPSAPLPLTSLAILLAPLSALIGCTL 180  
DB 121 ALSGCTLAISVPDGTALALTLPLGGLVPAITLPSAPLPLTSLAILLAPLSALIGCTL 180  
QY 181 SQSGDLGVISTPQLSHPIIGDWSIQVQVNDQRYYSQVSEYVLPKFEVTLTQPLVCS 240  
DB 181 SQSGDLGVISTPQLSHPIIGDWSIQVQVNDQRYYSQVSEYVLPKFEVTLTQPLVCS 240  
QY 181 SGSGALGVISTPQLSHPIIGDWSIQVQVNDQRYYSQVSEYVLPKFEVTLTQPLVCS 240  
DB 181 SGSGALGVISTPQLSHPIIGDWSIQVQVNDQRYYSQVSEYVLPKFEVTLTQPLVCS 240  
QY 241 MNSHGLNLTAKTYGKPKVQDVTLPPLPSFWGKKKNTKTEKINSANFSFNDEMK 300  
DB 241 MNSHGLNLTAKTYGKPKVQDVTLPPLPSFWGKKKNTKTEKINSANFSFNDEMK 300  
QY 241 MASHLAGTITALTTLTGTPVGAATLPLPLSPGILLALITLTLULINGSAPSRAGML 300  
DB 241 MASHLAGTITALTTLTGTPVGAATLPLPLSPGILLALITLTLULINGSAPSRAGML 300  
QY 301 NVMDSSNGLSLEYLDLSSPGEVILITTVESVTGISRNVSINVPFKQHDYIIEFPDYTVL 360  
DB 301 NVMDSSNGLSLEYLDLSSPGEVILITTVESVTGISRNVSINVPFKQHDYIIEFPDYTVL 360  
QY 301 AVMASSAGLSGTLALSSPGVGIITTVGVTGISAAVSTAVPLGHAITTIIGPATITVL 360  
DB 301 AVMASSAGLSGTLALSSPGVGIITTVGVTGISAAVSTAVPLGHAITTIIGPATITVL 360  
QY 361 KPSLNFATATVATADGQQLLEBRNNVITVQANVTWSSNSGNQMEAVQKINY 420  
DB 361 KPSLNFATATVATADGQQLLEBRNNVITVQANVTWSSNSGNQMEAVQKINY 420  
QY 361 LPSLIATATVATVTAAGGLTLGAAAAAVITVGAATTTGSSASGAGLMAVGLIAT 420  
DB 361 LPSLIATATVATVTAAGGLTLGAAAAAVITVGAATTTGSSASGAGLMAVGLIAT 420

QY 421 TVPQSGTFKIEPPILEDSSSELQKAVFLGSKSSMAVHSLFKSPSKYITQKTDENIKVG 480  
DB 421 TVPQSGTFKIEPPILEDSSSELQKAVFLGSKSSMAVHSLFKSPSKYITQKTDENIKVG 480  
QY 481 SPPELVVSGNKRILKEISYMWVSROGLAVAGKQNSTWMSLTPKNSMPPKACVIVYIEDDG 540  
DB 481 SPPELVVSGNKRILKEISYMWVSROGLAVAGKQNSTWMSLTPKNSMPPKACVIVYIEDDG 540  
QY 541 EISDVLKIPVQVFNKIKLWMSKYAEPSEKYSRLISVTPQDSIVGIVAVDKSVNLNM 600  
DB 541 EISDVLKIPVQVFNKIKLWMSKYAEPSEKYSRLISVTPQDSIVGIVAVDKSVNLNM 600  
QY 541 GIIISAVLLIPVGLPPLAILLITLTSIVLAGSGVLSLAVSTGPASIVGIVAVLSVALMA 600  
DB 541 GIIISAVLLIPVGLPPLAILLITLTSIVLAGSGVLSLAVSTGPASIVGIVAVLSVALMA 600  
QY 601 ASNDITMENVHLELNTGYIYGMFNPSPAVPOEGGLVYLDANLTQDYIDGVYDNEY 660  
DB 601 ASNDITMENVHLELNTGYIYGMFNPSPAVPOEGGLVYLDANLTQDYIDGVYDNEY 660  
QY 661 ASNAITWGAHVHGLTATGTTLGEMMASPAVPGGGLTVLTAAALLTATAGVTAAGT 660  
DB 661 ASNAITWGAHVHGLTATGTTLGEMMASPAVPGGGLTVLTAAALLTATAGVTAAGT 660  
QY 661 AERFMENEGHIVDIDHDSIGSSPHVKHPEFTWIMDTMNGRIYQEPFVTPDSTSW 720  
DB 661 AERFMENEGHIVDIDHDSIGSSPHVKHPEFTWIMDTMNGRIYQEPFVTPDSTSW 720  
QY 661 AGAPMGAGGHIVAHAPSLGSSPHVALHPGTTITLATIMGPAITGPGCVTPASTIST 720  
DB 661 AGAPMGAGGHIVAHAPSLGSSPHVALHPGTTITLATIMGPAITGPGCVTPASTIST 720  
QY 721 VATGFVISEDGLGLTTPPELQAFQFIFLNLPSVIRGEFPALETTFNYLKDATEV 780  
DB 721 VATGFVISEDGLGLTTPPELQAFQFIFLNLPSVIRGEFPALETTFNYLKDATEV 780  
QY 721 VATGPVISAALGGLTTPPELQAFQFIFLNLPSVIRGEFPALETTFNYLKDATEV 780  
DB 721 VATGPVISAALGGLTTPPELQAFQFIFLNLPSVIRGEFPALETTFNYLKDATEV 780  
QY 781 KVIIEKSKDFDIIMTSSEINATCHOQTLVPSBDGATVFPPIPTHGEPITVTALSPT 840  
DB 781 KVIIEKSKDFDIIMTSSEINATCHOQTLVPSBDGATVFPPIPTHGEPITVTALSPT 840  
QY 841 LVIIIGSLALPAILMITSAGIATAGGGLLVPSGAGATVLPPIPTHGEPITVTALSPT 840  
DB 841 LVIIIGSLALPAILMITSAGIATAGGGLLVPSGAGATVLPPIPTHGEPITVTALSPT 840  
QY 841 ASNAVOMILVKAEGEKSYSOSIILDTMNRQSLTKLSFSPNTVGSRRVQITAI 900  
DB 841 ASNAVOMILVKAEGEKSYSOSIILDTMNRQSLTKLSFSPNTVGSRRVQITAI 900  
QY 841 ASNAVOMILVKAEGEKSYSOSIILDTMNRQSLTKLSFSPNTVGSRRVQITAI 900  
DB 841 ASNAVOMILVKAEGEKSYSOSIILDTMNRQSLTKLSFSPNTVGSRRVQITAI 900  
QY 901 GVLVGSINGLALILIMPYCGEONMINFAPNIYIIDLTKKQJLTDNLKEKALFEWROG 960  
DB 901 GVLVGSINGLALILIMPYCGEONMINFAPNIYIIDLTKKQJLTDNLKEKALFEWROG 960  
QY 901 GVLVGSINGLALILIMPYCGEONMINFAPNIYIIDLTKKQJLTDNLKEKALFEWROG 960  
DB 901 GVLVGSINGLALILIMPYCGEONMINFAPNIYIIDLTKKQJLTDNLKEKALFEWROG 960  
QY 961 YORELLYQREDGSAFAGNYDPSGSTWSAFVLRCPLEADPYIDQNVLHRTYTWLKGH 1020  
DB 961 YORELLYQREDGSAFAGNYDPSGSTWSAFVLRCPLEADPYIDQNVLHRTYTWLKGH 1020  
QY 961 YORELLYQREDGSAFAGNYDPSGSTWSAFVLRCPLEADPYIDQNVLHRTYTWLKGH 1020  
DB 961 YORELLYQREDGSAFAGNYDPSGSTWSAFVLRCPLEADPYIDQNVLHRTYTWLKGH 1020  
QY 1021 QKNGEFWDPGRVYHSELQGNKSPVTLTAVITSLIGYKQVNPNDVQSHIFLSEFS 1080  
DB 1021 QKNGEFWDPGRVYHSELQGNKSPVTLTAVITSLIGYKQVNPNDVQSHIFLSEFS 1080  
QY 1021 GLSAGGPTAGAVIHSGGLGALSPLYTLTAVITSLIGYKQVNPNDVQSHIFLSEFS 1080  
DB 1021 GLSAGGPTAGAVIHSGGLGALSPLYTLTAVITSLIGYKQVNPNDVQSHIFLSEFS 1080  
QY 1081 RGISDNYTLLALITYALSSVSGPKAKELNMLTVRAEOGGMQFVSESKLSDSWQPSL 1140  
DB 1081 RGISDNYTLLALITYALSSVSGPKAKELNMLTVRAEOGGMQFVSESKLSDSWQPSL 1140  
QY 1081 AGISATTLALITLALSSVSGPKAKELNMLTVRAEOGGMQFVSESKLSDSWQPSL 1140  
DB 1081 AGISATTLALITLALSSVSGPKAKELNMLTVRAEOGGMQFVSESKLSDSWQPSL 1140  
QY 1141 DIEVAAYALLSHFLQOTSEGIPIKMWLSRORNSLGGPASTODTVALKALSEPALMNT 1200  
DB 1141 DIEVAAYALLSHFLQOTSEGIPIKMWLSRORNSLGGPASTODTVALKALSEPALMNT 1200  
QY 1141 AIGVAATALLSHFLQOTSEGIPIKMWLSRORNSLGGPASTODTVALKALSEPALMNT 1200  
DB 1141 AIGVAATALLSHFLQOTSEGIPIKMWLSRORNSLGGPASTODTVALKALSEPALMNT 1200  
QY 1201 ERTNIQVTVTPSSPSP-----LAVQPMANVISANGGPAICQNLNV 1243  
DB 1201 ERTNIQVTVTPSSPSP-----LAVQPMANVISANGGPAICQNLNV 1243  
QY 1201 GATTAIGTVTPSSPSP-----LAVQPMANVISANGGPAICQNLNV 1260  
DB 1201 GATTAIGTVTPSSPSP-----LAVQPMANVISANGGPAICQNLNV 1260  
QY 1244 YNFKAGSSRRRSIQNOEAFDIDVAVKENKIDLVNDVCTSFSGRSGMALMEVNL 1303  
DB 1244 YNFKAGSSRRRSIQNOEAFDIDVAVKENKIDLVNDVCTSFSGRSGMALMEVNL 1303  
QY 1261 TATLASSSSAAASIGAGAPALAVLALALALAVLAVLAVLAVLAVLAVLAVLAVLAVL 1320  
DB 1261 TATLASSSSAAASIGAGAPALAVLALALALAVLAVLAVLAVLAVLAVLAVLAVLAVL 1320  
QY 1304 LSGFVWPEASISLSETVKRYVDHGLNLYLDSVNETQPCVNIIPAVNFKVSTQDASVS 1363  
DB 1304 LSGFVWPEASISLSETVKRYVDHGLNLYLDSVNETQPCVNIIPAVNFKVSTQDASVS 1363  
QY 1321 LSGPWPVPSGISLSTGLVLTGTAHGLLALTLASVAGTGPCVAPPAVAPLVSATGASVS 1380  
DB 1321 LSGPWPVPSGISLSTGLVLTGTAHGLLALTLASVAGTGPCVAPPAVAPLVSATGASVS 1380  
QY 1364 IVUYIEPRKQAVSYNSEVYKLSGCDLCSVDQGRPREDEAGSGHHHSVIFICFLL 1421  
DB 1364 IVUYIEPRKQAVSYNSEVYKLSGCDLCSVDQGRPREDEAGSGHHHSVIFICFLL 1421  
QY 1381 IVATTGPAAGAVASTASGVLLSSCALCSAVGACPCGAGASGSHHSVIFICPPLL 1438  
DB 1381 IVATTGPAAGAVASTASGVLLSSCALCSAVGACPCGAGASGSHHSVIFICPPLL 1438

RESULT 15

AAB50676 standard; Protein; 1508 AA.

XX



**Tue Jan 20 08:47:22 2004**

**us-10-020-095-4.rag**

Page 18

D<sub>b</sub> 1356 CCNTRPBGKSMMLAEIDALSGRFDAGVHTLSIEDLQREVEHEKDPTKNAYFNPZLG 1415  
Q<sub>y</sub> 1340 TQFCVNIPAVANFKYSNTQDASVSIVDYEPBRQAVRSEYNEVLSSCDLCSVDQGRRP 1398  
Db 1416 RPYCLSLSDVTYYGADQKPANFRIVDYDDDEEQLKMITYAAKQIRSIQEKG--EDCMP 1472

Search completed: January 15, 2004, 18:12:20  
Job time : 59 secs

	Query Match	19.6%; Score 1441; DB 4;	Length 1474;
	Best Local Similarity	29.1%; Pred. No.5.9e-11;	
	Matches 451; Conservative	266; Mismatches 551;	Indels 284; Gaps 59
Qy	12 LLCVCTAALVAAPRPFLVTAPGIIIRPGANTIGVELLEHPCSPQYTVAAELIKTASNLTV	71	:
Db	15 LLVLPLPDASVSGKRQYVVLLPSSLHT-ETTEKGCVLISYNLETVTGASLSLEVAGNSL	73	:
Qy	72 -SVIEARG-VFEKSEFKTLTLPSLPLNSAD--IYLRLVGRTODELIFSNSTRLSFET	126	:
Db	74 FTDLAEANDVLHCAVF-----AVPKSSNNEEWFLLTVQVGNPTOE---FKKRTVMAYNA	124	:
Qy	127 KRISLVFIOTDALKYKPQOEKFERIVLTLESDEPKYKTSINLI-IKOPKSNLIQQOMIQSD	185	:
Db	125 EDSLVAFVQDTSLTKPGQATKFRVVSNDENHLPLVLYIQDPKNRIAQWOSFOLE	184	:
Qy	186 LGVISKTFOLSHPILIDWSIOVOVD--QTYOFSQVSEYVLPKEFVTLQTPLYCSMNS	243	:
Db	185 GGLQGQSFPLSBEFPQGSYKVNVQKKSGGRREH-PFTVESEVLLPRFEVQVTVPKITITLE	243	:
Qy	244 KHLNGITTAQTYTKPKPVKGDTLTFLP-----SFMGK	276	:
Db	244 EEMNVAISCGLYTYPGVGHATVASICRKYDSPADCHGEDSGAFCEKFGQLNSHQCFQO	303	:
Qy	277 KKNTTKTFKLINGSANFSFNDEEMGNWDSNGLSSEYLDLSGPPEILL-----TVTESV	331	:
Db	304 VK-TKVPQLKR-----KEYEMK-----LHTEAOIOEGIVVELTGQSSITTITI	347	:
Qy	332 TGISRNASTVNFQRKHDIYILEFDYTVLKPSLNFTATVKVTRADGNOLTEERRNNVI	391	:

```

Db      348 TKLS-FVKVDHSHFROG---IPFFG-----QVRLVDKGVPIP-----NKVI 384
Qy      392 TVTQRTNTEWMSGNSGNGKQKAEAVQ-KINTY--VPQSTFKIEBP-----IIEDS 438
Db      385 FI-NGENANYY--SNATTDHGLVQFSINTTNVMTSLTVRNVKDRSPCYGOWVSEH 441
Qy      439 SELQKAYFLGSKSMVAHSLFKSPSKTYIQLKTRDEIKVGS----PFLVVSNGK--R 492
Db      442 BEAHHTAY-----LVFSPSKSFVHLEPMSEHLPCHGHTQYQAHYIILANGTLG 489
Qy      493 LKELIS--YMWVSROQLVAVG-----KONST---MPSLPEMSWTPKACVIVYIEDDG 540
Db      490 LKKLSFYLLIMAKGAI VRTGTHGLLVQEDMKGHFSISIPKSDIAPARLLIYAVLPTG 549
Qy      541 EIIIDVLKIPQVLKPKIKLYMSKRYAESEKXSLRISVQPSIYGIYAVDVSVMN 600
Db      550 DVIDSAKYDVENCLANKVDLSFSPSOLPASHAHLRVTA--PQSVCLRAVDQSVLLMK 608
Qy      601 ASNDITMENNVHEL-ELVNTGY-----YLGMPMS-----FA 631
Db      609 PDAELSSASVYNLPEKDLTFPPRPLNDQDDEDCINRNVTYINGITTPVSTNEKMY 668
Qy      632 VFQECGLMWLTDAWLTKDYIDGVYDNE-----YAREMEENEGHIVDIDHESL 680
Db      669 FLEDMGKAFNTSKIRKQKCPQLOQYEMHGPBGLRVGFYSVDVMGRGHALVHEE--- 725
Qy      681 GSSHH--VRKHPEPTWIMLDTNMGYRYQEFVTVPSDTSWATGVISDGLGLTT 737
Db      726 ---PHTEVRKYPEPTWIMDLVAVNSAGVAVGTVDPTITWEMKAGFCSEDEGLIIS 782
Qy      738 TPVELQAFOPPEFLANLPYSVIRGESEFALBITTFNYLKDATEVAVYIEKSKDFILMTSS 797
Db      783 T-ASLRAPQPPFVELTMRPYSVIRGEAFLTKATVANTLPKCRVSVQLEASAFIAYVEK 841
Qy      798 E-----INATGHQOTLLVPSSEDAVTLFPIRTHLGEIPITVTLAS-----PTA 841
Db      842 EQAPHCICANGRO-----TVSWAVPKSLGVNFTVSALEASQELCTGEVPSV 890
Qy      842 -----SDAVTOMIIVKAGIEKYSQSILDLTNRLOSTIKTISFSPPTVTGSRVQ 896
Db      891 PEHRKQTVIKPLLVPEEGLEKETTFNSLLCPSGEVS---EELSLKLPVNVESARAS 947
Qy      897 ITAIGDVLGSPINGLASIRMPYCGGEONMNFANITYLIDYLTKKKQLTNLEKALSIF 956
Db      948 VSVUGDILGSMQNTQNLQMPYCGGEONMNFANITYLIDYLTQQLTPEIRISKALGY 1007
Qy      957 MRQGYORELLYQREBGSFSAFG-NYDPS-GSTWLSAFVLRCLFLEADPYIDIDQVNLHRTY 1014
Db      1008 LNTGYQQLNKHAYDGSYSTGERYGRNQNTWLTAFVLTKEFAQARAYIFIDEAHITQAL 1067
Qy      1015 TWLKGHQKSGNEFNDPGRAVHISELOGKNSPYTTATAYVLSLGRYKQPNIDQESIH 1074
Db      1068 TWLSQROKQDNCCFRSSGSLNNNAIKGVEDEVTLISAYITALLLEIPLVTHPVVRNALFC 1127
Qy      1075 LES-----EFSRGISDNVTLALITVALSSVGS--PKAKBALNMLTWABOEGCQMF--- 1124
Db      1128 LESAMKTAQEGDHG--SHVYTKALAVYAFALAGNODKREKVLKSLNEEYKXDSVHERP 1186
Qy      1125 VSSEKSLSDSWQPR--SLIDIEVAAYALSHFLQFQ---TSEGIQ---IMKWLSCRQNSL 1175
Db      1187 QKRPAPYGHFEPQAPSAEVEMTSYVLLA--YLTNQPAFTSBDLSANINIVMITKQONAQ 1245
Qy      1176 GGPASTODTVALKALSEFALMNTERTNIOVTYVGPSSPPLAVVQ----- 1222
Db      1246 GGFSTQDTVALALSKYGAATFT-RTGKAQVYTIQSSGFSSSKFYQDNRRNRLILQOVS 1304
Qy      1223 ---PMAVNISANFGFALICQLANVYVYKAGSSRRRSRSTIONQAFDLDVAVK---ENKD 1275
Db      1305 LPELIGEYSMKVTEGGCVYIQTSLKYN-----LPEKEFPFAPLGVQTLPTQCD 1353
Qy      1276 DLN-LHVDLNVCTSGSPG---RSGMALMEVNLISGFVWVPSAISL---SETVKKVEYDHG 1328

```

```

Db      1354 EPRAHTSFOISLSVSYTGRSASNMALVDYKWSGFIPLKPTVKMERSNHVSRTVESSN 1413
Qy      1329 KNLVYDVSVENTQPCVNIAPVNRFKVSNTOBASVSIYDVEYPPROAVRSYNS 1380
Db      1414 HVLITDKVSNQTLSTLFTVYLQDVPVRDLKPAIVKDYIYETDEFAIAEYNA 1465

RESULT 2
US-08-447-411-2
; Sequence 2, Application US/08447411
; Patent No. 5773243
; GENERAL INFORMATION:
; APPLICANT: FRITZINGER, DAVID C.
; APPLICANT: BREDEHORST, REINHARD
; APPLICANT: VOGEL, CARL-WILHELM
; TITLE OF INVENTION: DNA ENCODING COBRA C3, CVF1, AND CVF2
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/447,411
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/043,747
; FILING DATE: 07-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5773243man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 1126-101-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1651 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-447-411-2

Query Match      13.2%; Score 973.5; DB 1; Length 1651;
Best Local Similarity 23.5%; Pred. No. 9,8e-72;
Matches 391; Conservative 278; Mismatches 649; Indels 343; Gaps 61;

Qy      1 MGPRLITAAHLICVCTAALAVAPGPRFLVTAAGIIRPGANVTIGVEL-EHSPQVTVK 59
Db      1 MEGALYLVALLIGPPGS---SHGALYTLITPAVIRTDTEQDILVEAHGDSFPKSLDIF 57
Qy      60 A-----ELTKASNLVTVLEABGVFEKGFKTLTLPSPUNSADEIYELKVTGRTODEIL 115
Db      58 VHDPPRKQKTLFQSRVDMNAGSMFVT--PTIKVPAKELNKSQKQYVYVAVTGQVA 114
Qy      116 FSNSTRLSFETKRSIVFIQTDKALYKPKQEVKFRIVTLFSDFPKPYKTSNLILIKDPKSNL 175
Db      115 LKRVLLSYSG--PFVICTDKGIYTPGSPYRVRFVS--DHMHHRDKTVIYE----- 164
Qy      176 IQQWLSQSDLVYISKTFLQSSHPI-----LDWMSIQOVN---DOTTY 216
Db      165 -----FQTPREGIV-----VSSKPVNPSGSIIRPYNLPELVSGTAKVAVAKYEHSPESYT 213

```



217 OSFOYSEYVLPKPEVTLQTP---LYCSNNSKHLNGITTAKYKPKYKQVADVTLLTLELPSF 273  
 214 AFYDREYVLPKPEVTLQTP---LYCSNNSKHLNGITTAKYKPKYKQVADVTLLTLELPSF 271  
 274 WCKKXNI-----TKTEKINGSANPSPNDEEMKVMSSNGLSEYLDLSSGPEIILLTYTE 329  
 272 DDAKSIIPDLSLRIPIIDSGEATLKRDLSRFPQDLNOLVGHLLYVS-----VTVITE 325  
 330 SVTGISRVNSTVNFPEKHNDYIIEFPDYTVLKPISLNFATVATVTRADGNQLTLBERNNV 389  
 326 SSGDMVVEEGGIHIVTSPYQYIFTKPKYFKPGMYELTYVTPDSSPA-----H 378  
 330 VITVQKNTYEWSSGNSGNOMEAOKINTVP---OSGTFKIEPPIEDSSELOKAY 446  
 379 VPVVEALHSE-----GTLSDGTAKLILNTPLNIQS-----LPITVRTHGDLPRE 425  
 447 FLGSSSMVAHSLFKSPSKTYIQLKTRDNINYS--PELVVSGN-----KALKELSTNV 500  
 426 ROALISMTATVATQOGGSENYLHVAITSTEIKPGDNLVFNENVNGANSNLQIKFTYLI 485  
 501 VSRGOLVAVGK-----ONSTWPSLTPENSMTPKACVIVYIIEDEGEIISDVLKIPVQ-- 552  
 486 LNKGIIFKVGRRDQGLVYTMNLHITPDLIPSRFVAYIYOVGNNEIADSVWVDVKT 545  
 553 ---LVFK-----NKIKLYMSKVAPESEKVSLSRISTVQPDSTVIGIYAVDKSVNLMNS 602  
 546 CMGTLVVGKASSRDRIQ-----KPGAANKIKLE--GDPGARVGLVAVDKAVVYVYNDK 596  
 603 NDITMENVHELELNTGYLGMFMSFVFOECGLMVLTDANLT----- 647  
 597 YKISQAKIWDITEKSDFGCTAGSGONNLGVFEDAGIALTSTNLTMTKORSAKCPQANR 656  
 648 -----KDYIDGYDN-----AEYEFMEENG----- 670  
 657 RRRSSVLLDSTAKAQAQDQGLAKCCGDMHNPMTGCKRKYIOEGACRAAFIE 716  
 671 --HIV-----DIHPSLGSPPHV-RKHPEETIML-----DTWNG 702  
 717 CCHYKGIARDENQRESEFLARSDPEDELFGDDNIISRDPSMWLMBELTGEBNNG 776  
 703 YRIYQEFYVTP---DSTTSWANGFVISEDLGLGTTT-----PYELOAFQPPFIE 751  
 777 IS-----SKTVPFYLRDSITTW-----ELLAVGLSPTKGIQVAEPYEITVMQDFID 823  
 752 LNLPSVIRGEEFALETITFNLYKDAVEKVIIEKSDKEDIMTSSEINATHQDTLVP 811  
 824 LRLPSVAVKNEOVEIRAILYNTADEDIYVELIYNPAF---CSASTEGORYRQOPFK 879  
 812 SEDGATVLPFIRPTHLG--EIPITVATASPTASDAVQWILVKAAGIEKSYQSILDLT 869  
 880 ALSSRAVPFVILPBGJLHDEVIASVREGELASDGRKKLVPPGERKNITITIELDRS 939  
 870 DNRLOST--LKTLSGSPPNVTYGSERVOITAGD---VLQPSING--LASLIMPYGC 921  
 940 VKGVGTOELTVIAKLDKVPDTEVEETISVIGDPVAQIIEHSIDGSKLMLHITTPSGC 999  
 922 GEONMINIFAPNI---YILDYLTKKKQJDNLKEKALSFRQCYQRELLYQREDGSEFARG 978  
 1000 GEONNITMTSPVIATYLDATQWELGVDRTEAIKQMTIYAQOMVYKKAHDHYAALT 1059  
 979 NYDPSGSTWLSAFVLRCELEADPYI-DIDQNVLAHRTYWL--KGHOKSNGEPMDFGRVHS 1036  
 1060 N--RASSSWLTVAVVKVLAWSNWKDISHEIICGVVKMLINRQOPDVFKEANPVIHG 1117  
 1037 ELQGNKS---PVTLFATVITSLGGRKYQPNIDVOESIHLESEFSRG-----I 1083  
 1118 EMLGGTGAEPASLTAFTVITALLBSRV-----CQEQINIIDSINKATDYLKKYEKL 1172  
 1084 SNNTYALITVYLSVSGPKAKEALNMLTWRAQEGCGQFWVSSEKLSDSQPSLDIE 1143  
 1173 QRPYTTALVAVALA-----ADDRINDDRVLAASGTGRW-----EENYARHTNIE 1218  
 1144 VAAVYLLSHFLQFOTSEGIPIRWKLSRQNSLSGFGFASQDQTTVALKALSEFALMNTER- 1202

1219 GTSVALLLMLKKKAEVGPVVRMLIDQKYGGTYGTOATVWFOALAEYEIQPTHQD 1278  
 1203 TNOVTGSPSSPS-----LAVQPMAY-----NISANGFPAIQOLNVAVVK 1247  
 1279 LMLDSTIKLPEREVDERSINDRNAVQARTVETKLNEDFTVASASDGRATMTITLVYNAQ 1338  
 1248 ASGSSRRRSRSLONCEAPFLDVAV-----KENKDLNHDVLANVTSFSGPSGSMALME 1300  
 1339 L--REDANVCNK--FLDVDSVENVELNMLQAKGKXALRLKICTRYLIGVDSMTMTIID 1392  
 1301 VNLSGFVWPSEAL-----SLSEYKAYVEYDH-----GKLNLYDSVNETQ--FCVNIIPAV 1349  
 1393 ISMLTGFPEDDELKRLSNGVDYRYSKFEIINNMAQKGTVAIYLDKVSHEDECHAFKIH 1452  
 1350 RNFKYSNTODASVSLVDYPERRQAVRSYNSSEVYKLSDDLCSDVQCGPCEDGASGSHH 1409  
 1453 KHFEVGTQPSGVKAYSYNLDEOCTKPYHDKETGLINKICHGNICCAEBTCSLMOQ 1512  
 1410 -----SSVIFIFCFKL-----YEMEL 1426  
 1513 KKIDQLRIQKACQANVDVYVYKTKLRLIEKDGNDIYFMDV 1553

RESULT 3  
 US-08-793-126-1  
 ; Sequence 1, Application us/08793126  
 ; Patent No. 5849297  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Harrison, Richard Alexander  
 ; APPLICANT: Fairies, Charles Timothy  
 ; TITLE OF INVENTION: MODIFIED HUMAN C3 PROTEINS  
 ; NUMBER OF SEQUENCES: 2  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: HALE AND DORR LLP  
 ; STREET: 60 State Street  
 ; CITY: Boston  
 ; STATE: MA  
 ; COUNTRY: United States of America  
 ; ZIP: 02109  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent in Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/793.126  
 ; FILING DATE: 07-FEB-1997  
 ; CLASSIFICATION: 424  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Baker, Hollie L.  
 ; REGISTRATION NUMBER: 31,321  
 ; REFERENCE/DOCKET NUMBER: 102286.377  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (617) 526-6000  
 ; TELEFAX: (617) 526-5000  
 ; INFORMATION FOR SEQ ID NO. 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1663 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-793-126-1  
 Query Match 12.7%; Score 931; DB 2; Length 1663;  
 Best Local Similarity 23.0%; Pred. No. 3.6e-68;  
 Matches 376; Conservative 286; Mismatches 615; Indels 356; Gaps 62;  
 3 GPPITAAHLICVCTALAVAPGPFVTAFCIIR-----PGANNITGVELL 49  
 6 GPSLL-----LILLTHPLALG-SPWYSIITBNILIESEETVLEAHDAQDVPVTVVH 60

```

QY 50 EH-----CPBQVTVKAEKLLKTASNLTVSV-LEAEVGFEEKSGFETTLTPLSLNSADEIY 102
D 61 DPECKLVLSSEKTVLTPATNMGNTVFTIIPANEFPKSGRKNFVTVQA---TFGTVV 117
QY 103 ELARTGTODEILFNSNTRLSFETKRISVPIQTKALKYKQEKFIIVTLFSPFKRYKT 162
D 118 EKVVLVLSQSGYL-----FLOTDKTYTTPGSTYLYRIFVTHNKLTVGR 161
QY 163 SLNLLIDPKSNLIQO--WLSQSDLGVIKTFOLSHPIIDGMSIQOVND---QTYQS 218
D 162 TWMNINENPEGIPVKQDLSLSONQGLVPLSMJPELVNMGQMKIRAYENSPQOVSTE 221
QY 219 FOVSERYVLPKEVTLQ--TPLYCSMNSKHLNGITTAITYGKPKYGDVTLTF----- 268
D 222 FEVKEVYVLPSEVIVEPEKFFYYIYNEKGLVITITAFVLGKVEGAFVIFGIQDEQR 261
QY 269 -----LPLSWGKKKNTTKTFKINGSANFSPNDEKMNWDSNGSLSEYIDLSSPG 319
D 282 ISLESUKRIPIEDSGEVLRSKVLDDGVONPRAEDLVGKSLVSA----- 328
QY 320 PVELTIVTESVTGISRNVSTNVFFKQHDYIIEFFDYTVLKPSLNTFATVYKVRADGN- 378
D 329 -----TIIHSGSMVQERSGPIYVSPYQIHTTKPKFKPMFPLDMFVTPNPGSP 383
QY 379 --QULEBERNNVITYTORNYTEWGSNGSNOKMEAVOKINTVPOSGTFKLEPIL 436
D 384 AYRPVAVOGEDTVYQSLQDGVAKLSINTHPQK-----PLSTIVRRKQELS 432
QY 437 DSSLOLKAFLGSKSMVAHSLFKSPSKYIQLKT-----RDENIKVSGPEFLVVSNG 490
D 433 EABD-----ATRTVQALPYSTVSGNSNNYLSVLRTELPEGTINVLMLMDRAHE 484
QY 491 KRLKELSYMVVSRQOLVAVGK-----ONSTMFSLTPENSWTPKACIVVY---IEDDG 540
D 485 AKIYTYTYLLNKKRLKAKGRQVREPGDVLVPLSTTDFIPBFLVANYTYLLIGASQR 544
QY 541 EISDVLKIPVQ-----LVFKNKIKLYMSKVKA-----PSEKVSRLISVTOPSIVGI 589
D 545 EVVADSVWVVDVQDSVCSLVVKG-----QSEDRQVPVQCGQMTLKIBGDHARV-L 595
QY 590 VAVDKSVNLMAASDITMENVHELELYNGYIYGMFNSFAYVQEGCLWVLTDA----- 644
D 596 VAVDKGVFLNKKKLTOSKIMDVEKADIGCTPGSKDAGVFSDAGLTFSSSGQOTA 655
QY 645 -----NLTDYIDGVVDN-----AEXAERFME 666
D 656 QRAELQCPAPARRRRSVOLTTERKMDKVKYKPKELRKCCEDEKMEKMPRSCQRTFFIS 715
QY 667 ENB-----GHIVDIH-----DPSLGSSPHV--RKHPEPTWIM- 696
D 716 LGBACKKVPLOCCNVITELRQGHARASHLGLARSLNDEBIIAENIYSRSEFPESWLMV 775
QY 697 -----LDNMGRIYIOEFVTVPDSITSVNATGFIASEDLGGLTTPELOAFQ 746
D 776 EDLKEPPKNGISTKL-----NMNIFKDSITWEILLAVMSDKKGI-CVADPEVTVMQ 827
QY 747 PEFIFLPLPVSVINGEPALEITIFENLKATEKVIIE--KSKKPIILMTSSSINATGH 804
D 828 DPFIDLPLPSVSVNEQVEIRAVLYNY-RONQELKAVVELLHNAFCSLATTKR---RH 882
QY 805 QOITLIVSEBDATVLPFIRP--THAGEIPITVTAISPTASDAVQMTLVKAGIEKYSQ 862
D 883 QOITIPKSLSLSPYIYVPLKTLQLEVENVAAYHHFISGVRKSLKLVPEGRBMKV 942
QY 863 SILIDLTDNRL--OSTLKTSFSFPNTV-----TSESR---VOITAGDVLPSPING- 910
D 943 AV-RTLPERLGRGVOKE---DIPPADLSDQVDPTESETRILLQGPVQMTEDAVDAE 998
QY 911 --LASTLIMPVGGCGGOMNINFAPNIYIIDLYTKKQULDNLKER---ALSTMROGYQBEL 966
D 999 RLKRLIYTPSGCGGONNIGMTPTVIAHYIYDETEQWERFGLEKQGLALEIKKYTQOLA 1058
QY 967 YOREDSGFSFAFGNYPGSGTWLSAFVLRCPLEADPYIDIDQNVLHRTYWL-KGHQSGNG 1025

```

```

D 1059 FROPSSAFAFVRAAP--STWLTAAVYVVKFSLAVNLIAIDQVLCGAVKMLILEKQKPDG 1116
QY 1026 EFWDPGRVHSELOG-----NKSPTVTLTAIYVTSILGYRKYQONIVD-OESII----- 1073
D 1117 VFOEDAPVTHQEBITGGLRNNNEKOMALTAFLVLSL-----QKADICEBQVSLGSGIT 1170
QY 1074 ----PSEFSRSGISDNVTLALITVALSSVSGSPKAKEALNMLTPRAEOBGMQFVWSSES 1129
D 1171 KADDFLEANY--WMLQSGSYVALAGVLAQMGRKGLKFL-----TA 1214
QY 1130 KLSDSWQ--PRSLDIEVAAYALLSHFLOPQTSSEGI-PIRWLISROPNLSIGFPASTQDIT 1185
D 1215 KDKRMEDRQKOLYNTAETSALLA-LLOLKOPDFVPVVRWMEORRYGGGYGTOATF 1273
QY 1186 VALKALSEFAA-LMTERNTIOQTVVGPSPSPLA-----VYQPMVANS 1229
D 1274 MVFQALAQYQKAPDHOELNDVSLQPSRSSKITHIRIMESASLLRSEBTEKNEGFTYV 1333
QY 1230 ANGFGEAICQLNVVNVKASGSSRRRSIONGEAPDLDAV-----KENKDLNHVDL 1282
D 1334 ABEKGQGLSVVTMTHAKAKQDLTKN-----FDLKVTIKAPETEKRPQDAKNTMIL 1386
QY 1283 NVCTSGSGRSGMALMEVNLISGFVPSB-----AISLSEYVKVY-----DHGKLN 1332
D 1387 EICTRYRGQDAMTSLIDISMMTGFAPDITDLKQLANGVDRYISKYELDKAFSDRYTLII 1446
QY 1333 YLDSVNETO-FCNINAVNFKVSNQODASVSVVDYEPERRQAVRSENVSE--KLSQSD 1388
D 1447 YLKVSHSEBDCIAFVHOIFVNELIOPAVKAYAYNLEBSCTRFYHPEKEDKLNK-- 1504
QY 1389 LQSDVQCRPCED 1401
D 1505 LCRD-BLCRCABE 1516

```

```

RESULT 4
US-09-132-271-1
; Sequence 1, Application US/09132271
; Patent No. 6221657
; GENERAL INFORMATION:
; APPLICANT: Harrison, Richard Alexander
; APPLICANT: Farries, Charles Timothy
; TITLE OF INVENTION: MODIFIED HUMAN C3 PROTEINS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HALE AND DORR LLP
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: United States of America
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/132,271
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/793,126
; FILING DATE: 07-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Baker, Hollie L.
; REGISTRATION NUMBER: 31,321
; REFERENCE/DOCKET NUMBER: 102286,377
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 526-6000
; TELEFAX: (617) 526-5000
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:

```

LENGTH: 1663 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-132-271-1

Query Match 12.7% Score 931; DB 3; Length 1663;  
Best Local Similarity 23.0%; Pred. No. 3,6e-68;  
Matches 376; Conservative 286; Mismatches 615; Indels 356; Gaps 62;

3 GEPILTAHLICVCTALAVAPRRLVAPGIR-----FGAVTIGVELL 49  
6 GSSL-----LILHLPLALG-SPWYSITTPNILLRESEETWLEAHDAQGVPLVTVH 60  
50 EH-----CPQVYAKELTKTASNLTVSV-LEAEVFEKSPFKLLTSLPLNSADELY 102  
61 DPPGKLVLSSEKTVLTPATNMGNVTFIPANREFSEKGRKFTVA-----TEGTQV 117  
103 ELRVGTQDELLFNSSTRLSFETKRISVFIQTDKALYKQEVKFRITVLFSDFKPYKT 162  
118 EKVLVLSQSGYL-----FIQTDKITPGSTVLKRIFTVNHKLIPGR 161  
163 SLNLIKDKPKSNLIQO-WLSQSDLGIVSKTQOLSHPILGWSIQOVND---QTYQS 218  
162 TWMVNIENDEGI PVKQDSLSQNOGLVLPISWDIPELVNMQWKIRAYENSPOQVFSTE 221  
219 FQVSEVLPKFEVTLQ--TFLYCSMNSKHLNGTITAKTYGKPVAGDVLTF----- 268  
222 FEVKEVLPKFEVTVLPEKFIYIYNEKLEVTITARFLYKQVETARFVIGIODEGR 281  
269 -----LPLSPGKKKNTTKFKINGSANFSFNDEMKVNMSSNGLSEYLDLSPG 319  
282 ISLPSIKRIPLEDSDGEVLSRKVLIDGVQNPRAEDLVGKSLYSA----- 328  
320 PVEILITTESVATGSRNVSTNVFKQHDYIIEFDYTVLKPISLNFATVATVADGN- 378  
329 -----TVLIHSGSDMWQARSGIPVTSFYQIHFTKPKFKPMFDMVFTVMDGSP 383  
379 --QTLLERRNNVITVQANTETWSSNSGNOCKEAVOKINVTYPOSTKIEPILE 436  
384 AYRPAVAVQGEDVQSLTQGDVAKLSINTHPSOK-----PLSITRYTKQELS 432  
437 DSEELQKAYFLGSKSMVAHSLFKSPSKTYIQLKT-----RDENIKGSPFELVNSN 490  
433 EAEQ-----ATRMQALPYSTVGSNSNNYHLSTRLKRLPGETLVNPLIRMDRAE 484  
491 KKLKELSYNVVSRGOLVAVGK-----QNSTWFSILTPENSWTPKACVIVY---IEDDG 540  
485 AKIRYTYVLIIMKGRLLKAGROVREPGQDLVLPISITTDPIPSFLVAVYTLIGASGR 544  
541 ELISDVCLKPVQ-----LVFKKIKLKYKVAE-----PEKVSLSRISTVQPSIVGI 589  
545 EVVAASVWVDVDSQVSGSLVYKSG-----QSEBRQVPCQOQTLKIEGDHGAHV-L 595  
590 VAVDSVNLMAANSNDITMENVVHELELVNTGYVYLGFMFMSFAVPOCGJMWLTD----- 644  
596 VAVDGVAVLANKKAKLTQSKIMDVVEKADIGCTPGSGKQYAGVFSAGLFTTSSGQOQA 655  
645 -----NLTKQYIDVYND-----AEYARFME 666  
656 QRAELQCPARARRRSVQLTEKMDKGVKPKELKCCEDGMRNEMPAFSCQRRTRFS 715  
667 ENE-----GHIVDIH-----DPSLGSSPHV--RKHPETWIM-- 696  
716 LGEACKVFLDCCNVITELRQHARASHGLARSNDEDIIEAENIVSRSEPESLMNV 775  
697 -----LPTNMGVRIYQEFVTVPSITSWATGVISEDGLGITTVPVLAQAO 746  
776 EDLKEPKKGIKTL-----MNIPLKDSITTWELAVSMDSKGI-CVADPFETVWQ 827  
747 PFIPLNLPYSVIRGEFALETITFNVLADATEVKVIL--KSDKEDILMTSSINATGH 804

828 DFDILRLPYSVVRNEOVEIRAVLVNY-RQNEKLVRELLHNPAFCSLATTKR----RH 882  
805 QOTLLVPSBDGATVLPPIRP--THLGEIPITVTALSPVADAVTMILVKAIGIEKYSQ 862  
883 QOTITIPKSSLSVYVLPKLTGQEVKRAVYHHNHSIDGVKSLKVPBGIMNKTV 942  
863 SLIDLDTNRL--QSTLKLSPSPPTNV-----TGSER---VOITAGVLAGPSING- 910  
943 AV-RTLDERLGRGEGVQ---DIPRALSDQVPTESETRILLGTPVACMTEDAVDAE 998  
911 -LASLRMPYCGGEGNMNFAFNITLIDYLTKKKOLTNLK---AISFMQGIQRELL 966  
999 RLKHLIVPSSGGEONMGMTPTVAVHYLBTEBQEKRGLEKQAGALELKKGYTOOLA 1058  
967 YOREDSFSAFGENVDPGSTWLSAFVLCPLEADPYIDQNVHRTTWL-KGHOKNG 1025  
1059 FRQSSAPRAAFVYKAP--STWLTAVVAVFSLAVNLAIIDQVLCGAVKWLILEKQPDG 1116  
1026 EFWDPGRVYHSELQGG---NKSPTLVAVYVTSLSGRYKQPNIDV-QEESI----- 1073  
1117 VQEDAPVTHQEMIGGLANNEXOMALTAFLVLS-----QEAIDICEOVNSLPSIT 1170  
1074 -----FLESFBSRGISDNTTALITVYALSSVSPKAKALNMLTWRAEQGGKQFVSSBS 1129  
1171 KAGDFLEANY-MNLIQSYTVAIAGYALAQMGRLLKGPLINKFLT-----TA 1214  
1130 KLSDSMO---PRSLDIEVAAYALLSHFOPTQSEGI-TIMWLSQRNSLGFASFTOFT 1185  
1215 KQKNWEDEPGKQLVNVEATSYALLA-LIQLDQDFVPVAVWLNQRYYGGYGSTQATF 1273  
1186 VALKALISEFAA-LMTERKTNIQVTVGPSSPSPLA-----VQPAVANVS 1229  
1274 MVFQALAQYQADAPHOGLINDVSLQPSRSKITHRIHWSASILRESEETENGFTYT 1333  
1220 ANGFQFALCOLNVVNVKASSSRRRSIQNEAFDLDAV-----KENKDLNHVL 1282  
1334 AEGKQGTSLVTVTMHAKAKQLTCNK-----FDLKVTIKPAPETKRPQDAKNTML 1386  
1283 NVCTSPSGRSGMALMEVNLISGRMPSF-----AISLSEVYKVEY-----DHGKNTL 1332  
1387 ELCTRYRDQDATWSLIDISMTWGPAPDTDLKOLANCVDRYISYELDKAFSDNNTLIT 1446  
1333 YLDSVNETQ-PCVNI-PAVRNFKVSNTPDASVIVDYEPFRQAVSYNSEV---KLSGCD 1388  
1447 YLDKYSHEDEDLAKKHQYRNVELIQGAVKVAIYNLESCTRFTYHPEKEDGKLAN-- 1504  
1389 LQSDVQGRPCED 1401  
1505 LCRD-ELCRCAE 1516

RESULT 5  
US-09-142-334-22  
Sequence 22, Application US/09142334  
Patent No. 6268485  
GENERAL INFORMATION:  
APPLICANT: Fairies, Timothy C.  
APPLICANT: Harrison, Richard A.  
TITLE OF INVENTION: Down-Regulation Resistant C3 Convertase  
FILE REFERENCE: 4-30443/A/IMO/PCF  
CURRENT APPLICATION NUMBER: US/09/142,334  
CURRENT FILING DATE: 1999-04-15  
EARLIER APPLICATION NUMBER: PCT/GB97/00603  
NUMBER OF SEQ ID NOS: 35  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 22  
LENGTH: 1663  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-142-334-22  
Query Match 12.7% Score 931; DB 3; Length 1663;

Best Local Similarity 23.0%; Pred. No. 3.6e-68;  
Matches 376; Conservative 286; Mismatches 615; Indels 356; Gaps 62;

```

Oy 3 GPRLLTAHLACVCTAALAVAPGPRFLVATGRII-----PGNAVITIGELL 49
Db 6 GPRSL-----LILLTHPLALG-SPWYISITTNILHLESEETMVLBAHDAQSDVPTVTVH 60
Oy 50 EH-----CPSQVTVKELKLTASNLTVSV-LLEAGVPEKSGFKTLTLPILNSADEIY 102
Db 61 DFGPKLVLSSEKTVLTPATNMGVNFTIIPANREFSEKRNKFNVTQA---TFGTVV 117
Oy 103 ELRYTGRTQDEILFSNSTRLSFETKRISVFIQTDKALYKQOEYKFRIVTLFSPFKPKYT 162
Db 118 EKVVLVLSQSGYL-----FIQTDKITYTPGSTVLYRIFVYVNHLLPVGK 161
Oy 163 SLNLLIDPKSNLLIQ--WLSQSDLGVSITKTFQSLSHILGDMGIQGVND---QTYQS 218
Db 162 TVMNVNIEPBGIPYKQDLSQNOGLVPLSMWIPELVNMQMKIRAYENSPOQVSTE 221
Oy 219 FQVSEYVLPKFEVTLQ--TPLYCSMNSKHLNGITTAKYTGPKYGDVTLTF----- 268
Db 222 FEVEEYVLPSEFVIVEPEKFPYIYNEKGLVITITARLYGKXEGTAFVIFGIQDGEQR 281
Oy 269 -----LPLSFMGKKKNTTKTFKINGSANFSFNDEKMNWDSNGSLSEYIDLSPG 319
Db 282 ISLPESLKRIPIDEGSGEVLSRKVLLDGVQNPRAEDLVGKSLYSA----- 328
Oy 320 PVELITVTESVTVGISRVSTNVFFKQHDYIIEFDYTVLKPBLNFAATYKVRADGN- 378
Db 329 -----TVLHSGSDMVQAKERSGIPVTSYQIHFTKPKYFPMPELWVFNTPGSP 383
Oy 379 -QTLBERRNNAVTVTVQRYNTEYSGNSGNOKMEAVOKINTVPOSGTFKIEPFLIE 436
Db 384 AYRPVAVQSGEDTVQSLQSGVAKLSINTHPSQK-----PLSTIVRKQKELS 432
Oy 437 DSSELQKAVLYGSKSSMAVHSLFKSPSKTYIQKT-----RDENIKVSGPELVVSGN 490
Db 433 EAEQ-----AVRTMQALPYSTVSGNSNNYHLHSLVLTRELPGETLVNVLHMDRAHE 484
Oy 491 KRLKELSGVNVSRQOLAVGK-----ONSTMFSLTPENSWTPKACVIVY---IEDDG 540
Db 485 AKIRYTYLLNMKRLKAGQVREPGQDLVLPILSTTDPISFRLVAYTLLIGASQQR 544
Oy 541 EILSDVLKIPVQ-----LVFKNKIKLYWSKYKAE-----PSEKYSRLISTVQPSIYGI 589
Db 545 EVVADSVWVVDKDCVCSGLVYKSG-----QSEDRQPVGQOMTLKIEBDHARVY-L 595
Oy 590 VAVDKSVNLMAVNSDITMENVHLELYNTGYLGMFNNSPFAVQECGLWVLTDA----- 644
Db 596 VAVDKGVFVNLKKNKLTQSKIMDVEKADICTPGSGKDVGAVPSDAGLFTSSGGQTA 655
Oy 645 -----NLTKDYIDGVYDN-----AAEAERPME 666
Db 656 QRAELQCEPAPARRRSVQLTEKMDKVQKPKELRKCECEGBEMENPWFSCQRTFIS 715
Oy 667 ENB-----GHIVDIH-----PESLGSPPHV--RKHFPEWTW-- 696
Db 716 LGEACKKVFLDCCNYITELRQBARASHLGLARSNLDEDIJAENIYRSRFSFESMLNV 775
Oy 697 -----LDTNMGVRIYQEFVTVPDSITSWVATGTVISEDILGLTTPVELQAFQ 746
Db 776 EDLKEPPKNGISTYL-----NMIPKDSITTWELLAVNSDKKGI-CVADPREVIVM 827
Oy 747 PFFFLNLPYVINGEFALEITTFNYIKATERYVITE--KSKRFDLMTSSINATGH 804
Db 828 DFFIDRLPYSVVNEQVEIRAVLYNY--RQNEQLKRVETELHNPFCSLATTKR-----RH 882
Oy 805 QQTLLVSEDCATVLPFRP--THLGEIPITVTLSPASDAVQOMLIVKAGEIKESYSQ 862
Db 883 QQTITTPKSLSPYIVLPLKTLQGEVEVAAVYHHFISGVKSLKVVPEGIRANKTV 942
Oy 863 SILDLTDNRL--OSTLKTLSFSPPPNTV-----TGSER---VOITAGVVLGVSING- 910

```

```

Db 943 AV-RTLDPERLGBEGVQK---DIPPADSDQVDPTESETRILLQSTPVAQMTEDAVDAE 998
Oy 911 -LASLIRMPYGCCEOMNINFAPIYILDYLTKKKQOLDNLKEK---ALSPRQGYORELL 966
Db 999 RLKHLIYTPSGCEQNMIGMTPTVIAVHIYLDTEQNEKRELEBQCALIKKGYTQOLA 1058
Oy 967 YQREDSFSAFGNYDPBSGTSWLSAFVLRCELEADPYIIDQNVLHRTYTWL-KGHOKSNG 1025
Db 1059 FQPPSAFAFAFVKRAP--STWLTAVYVVKVPSLAVNLIAIDSVLCAVKKVLEIKQKPDG 1116
Oy 1026 EFMDDPGRVHISELQGG---NKSPTVLTAVIYVSLGKRYKQYNIDV-QSISH----- 1073
Db 1117 VFQEDAPVHQEMIGIGIRNNNEKDMALTAFLVLSL-----QAKOICEQVWLPGSTIT 1170
Oy 1074 -----FLESEFSGISDNYTLALTLYALSSVSGPRAKEALNMLTPRAEQEGMQFWVSES 1129
Db 1171 KAGDFLEANY--NMUQSYTALAGYALAOGRKLGPKLNFLE-----TA 1214
Oy 1130 KLSDSMQ---PSRLDIEVAAVYALLSHFLQPTSEGI-PIRWLSROKNSLGFSTQDIT 1185
Db 1215 KDKRMEDPKQOLYNYEATSYALLA-LLOLKDPDFVPVVRMLNEORYYGGSYGSTQATF 1273
Oy 1186 VALKALSEFAA-LMTERTVIQTVTGSPSPFLA-----VQPMANVIS 1229
Db 1274 MVFOALAQYOKADPDHQLNLDVSLQPSRSSKITRHHMESASLSRSEETKENEGFTVT 1333
Oy 1230 ANGEFALCOLNVYVNVKAGSSRRRSIQNEAFDLDVAN-----KENKODLNVLD 1282
Db 1334 AEGKQGTLSVMTYHAKKADQLTCNK-----FDUKTIYKAPETEKRPQAKMTML 1386
Oy 1283 NVCTSFSGPSRGMALMEVNLISGFVWPSB-----AISLETYKVEY-----DHGKML 1332
Db 1387 EICTRYRGDDATMSLIDISMGTGFAPDTDQLQLANGVDRYISKYELDKAFSDRMTLLI 1446
Oy 1333 YLDSVNETQ-FCVNIPTAVNFKVSNQODASVSLVDYIEPRQAVRYSNGV---KLSSCD 1388
Db 1447 YLDKVSHEBDDCLAFKHQYFVNLIQPGAVKYVAYVNLSECTRFYHPEKEDGKLNK-- 1504
Oy 1389 LCGSDVQCGRCPED 1401
Db 1505 LCRD-ELCRCAEE 1516

```

RESULT 6  
US-08-487-283A-2  
; Sequence No. 2, Application US/08487283A  
; Patent No. 6355245  
; GENERAL INFORMATION:  
; APPLICANT: Evans, Mark J.  
; APPLICANT: Matlis, Louis A.  
; APPLICANT: Mueller, Eileen Elliott  
; APPLICANT: Nye, Steven H.  
; APPLICANT: Rollins, Scott  
; APPLICANT: Rother, Russell P.  
; APPLICANT: Springhorn, Jeremy P.  
; APPLICANT: Squinto, Stephen P.  
; APPLICANT: Thomas, Thomas C.  
; APPLICANT: Wilkins, James A.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE TREATMENT  
; TITLE OF INVENTION: OF INFLAMMATORY DISEASES  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Seth A. Fidel  
; STREET: 25 Science Park (Alexion)  
; CITY: New Haven  
; STATE: Connecticut  
; COUNTRY: USA  
; ZIP: 06511  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch, 1.4mb storage  
; COMPUTER: Macintosh Cetrus 610  
; OPERATING SYSTEM: System 7  
; SOFTWARE: WordPerfect 3.0



Db	1454	QIKHGHTVLQNSIPSSDFLCRRPFIIFELFVGFSLPATFYVYEHKRDCKTMYST--	1511
QY	1383	KLSCDCLSDVQGCRPCEDGA	1403
Db	1512	-----SNIKIQVCEGAA	1524

## RESULT 7

US-08-662-227-2  
; Sequence 2, Application US/08662227

**GENERAL INFORMATION:**

; APPLICANT: VOGEL, CARL-WILHELM  
; BREDEHORST, REINHORST  
; ADDRESSEE: WOKK WICHTATI

APPLICANT: ROCK, MICHAEL  
APPLICANT: FRITZINGER, DAVID  
TITLE OF INVENTION: RECOMBINANT PROTEIN

CORRESPONDENCE ADDRESS:

ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
P.C.  
ADDRESSEE: P.C.

CITY: ARLING

STATE: VA  
COUNTRY: USA

```

; ZIP: 22202
;
; COMPUTER READABLE FORM:
; MEDICAL EVENT 11-11-84 14:15

```

```

;
; MEDIUM TYPE: floppy disk
;
; COMPUTER: IBM PC compatible
;
; OPERATING SYSTEM: PC-DOS
;

```

```

; OPERATING SYSTEM: PC-DOS/Ms-DOS
; SOFTWARE: PatentIn Release #1.0, Version 1.0
CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: US/0  
FILING DATE: 14-JUN-1996

CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:

```

; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
;

```

TELEPHONE: 703-413-3000  
TELECOMMUNICATION INFORMATION  
REFERENCE/DOCKET NUMBER:

TELEPHONE: 703-413-2222  
TELEFAX: 703-413-2222  
INFORMATION FOR SEO ID NO.

SEQUENCE CHARACTERISTICS:  
LENGTH: 1642 amino acids

```

;      TYPE: amino acid
;      STRANDEDNESS: single

```

```

; TOPOLOGY: linear
; MOLECULE TYPE: protein
;

```

05-08-662-227-2

Query match	12.46	score 894	DB 2	length 1842
Best Local Similarity	22.58	Pred. NO. 4.4e-65		
Marches 361	Conservative 205	Mismatches 672	Indels 276	Gaps 56

00	1A	AA1AY-----APGPRFTAPGTTBPGNTTGVELL--FHQPSQWYKA-----EILKTAS	67
10	201	CONCEPVALIVE	225/
20	301	HEBIMACUCIB	072/
30			INUCIB
40			270/
50			QGPS

Db 241 -ENFHSITARYLGEEVEG-VAFVLFGVKIDDAKSIIPDSLTRIPIDGDKATLKRD 298

Qy 299 MKNWDSNGISEYLDSSQPEVILTYYTESVTGISRNASTNVEFKQHDYIIEFDYTT 358

[illegible]

QY 359 VLKPSLNFATVVKVTRADGQQLTEERRNNVITVTQRNYTEYWSGSGNSGQKMEAVQKI 418

Db 353 YFKPGMPYEITVYTNPDGSPA-----AHVPVSEAF-----HSMGTTLSDGIATKL 398

419 NTVPQSTFKIEPILDDSELO,KAYFLGSSMAVHSLFSPSKTYIQLKTRDENIK 478

DB 399 IINIPLENAQ---SLEITVRNHHGDIIPREBQATKSMALAYQUGSSSNYLHVAITSIEIK 455

00	VGS--FRELVSIGN----	KKENLSDIMVVSCKDLAAVAG-----	ONSIMFSLIPENSMI	326
4/9	----	----	----	----
05	----	----	----	----
06	----	----	----	----
07	----	----	----	----
08	----	----	----	----
09	----	----	----	----
10	----	----	----	----
11	----	----	----	----
12	----	----	----	----
13	----	----	----	----
14	----	----	----	----
15	----	----	----	----
16	----	----	----	----
17	----	----	----	----
18	----	----	----	----
19	----	----	----	----
20	----	----	----	----
21	----	----	----	----
22	----	----	----	----
23	----	----	----	----
24	----	----	----	----
25	----	----	----	----
26	----	----	----	----
27	----	----	----	----
28	----	----	----	----
29	----	----	----	----
30	----	----	----	----
31	----	----	----	----
32	----	----	----	----
33	----	----	----	----
34	----	----	----	----
35	----	----	----	----
36	----	----	----	----
37	----	----	----	----
38	----	----	----	----
39	----	----	----	----
40	----	----	----	----
41	----	----	----	----
42	----	----	----	----
43	----	----	----	----
44	----	----	----	----
45	----	----	----	----
46	----	----	----	----
47	----	----	----	----
48	----	----	----	----
49	----	----	----	----
50	----	----	----	----
51	----	----	----	----
52	----	----	----	----
53	----	----	----	----
54	----	----	----	----
55	----	----	----	----
56	----	----	----	----
57	----	----	----	----
58	----	----	----	----
59	----	----	----	----
60	----	----	----	----
61	----	----	----	----
62	----	----	----	----
63	----	----	----	----
64	----	----	----	----
65	----	----	----	----
66	----	----	----	----
67	----	----	----	----
68	----	----	----	----
69	----	----	----	----
70	----	----	----	----
71	----	----	----	----
72	----	----	----	----
73	----	----	----	----
74	----	----	----	----
75	----	----	----	----
76	----	----	----	----
77	----	----	----	----
78	----	----	----	----
79	----	----	----	----
80	----	----	----	----
81	----	----	----	----
82	----	----	----	----
83	----	----	----	----
84	----	----	----	----
85	----	----	----	----
86	----	----	----	----
87	----	----	----	----
88	----	----	----	----
89	----	----	----	----
90	----	----	----	----
91	----	----	----	----
92	----	----	----	----
93	----	----	----	----
94	----	----	----	----
95	----	----	----	----
96	----	----	----	----
97	----	----	----	----
98	----	----	----	----
99	----	----	----	----
00	----	----	----	----

00 527 PRACYIYYIEDDGEIISDVKIPVO-----LVFK--NKIKLYMSKVKAEBSEKVSIRI 578

Db 516 PSFRVAYIQVGNNEIVADSVWVDYKDTCMGLTVKGDNIQM-----PCAAMKIKL 567

QY 579 SYTQPDSTVGIVAVDKSVNIMNASNDITMENVHELINTGYILGMFNNSFAVQEGCL 638

Db

568 E-GDPGARVGLVAVDKAVYVLNDKKYSIQAKIWDLTLEKSDPGCTAGSGQNNLVGFEDAGL 626

639 WILTDANI-TKD----- 649

Db 627 ATTSTNTLKÖRSAAKCPQANRRRRSSVLLDSDASKAAEFQDDLRKCCEDVNHENP 686

QY 650 -----YIDGVYD-NAEYARFM--EENE-CHIVDIDHF 678  
| | | | | : | | : | |  
| | | | | : | | : | |

Db 687 MGYTCEKRAKYIQEGDACKAAFLCCECRYIKGVNDNQSESELF LARDNEDGFIDSDII 746

679 SLGSSPHVRKHFFPEIWIWL-----DTMGIYRIYDFEVIWIPDSIISWATGFISSDLG 732

Db 747 S-----RSDPFKSWLWLT KDLT EEPNSQGISSKTNMSFYLRDSITTWVLA VSPFPTKG 799

QY 1233 PEPALCOLNVTYVNVKASGSSRRRSTIQNEAPDLVAKENKDILNHD-----LAWCT 1286  
 Db 1318 DEKATMTITLTFYNAQ-----QEKANVCN- PHILANVS- NINHMANAKGALMLKICT 1369  
 QY 1287 SFGSGRSGMALMEVNLISGFVNPSPAIS-----LSEYKAYEYDHGKLN-----LYLDS 1336  
 Db 1370 RILGEVDSTMTIIDSMLTGLPDADDLTRLSGVDRIYSRYEVNNMAQKVAIILYLNK 1429  
 QY 1337 VNETQ-FCVNIPIAVNFKVSTQDASVSIVDYEPBRCAVRSYNSVUKLSGDLCSDVGC 1395  
 Db 1430 VHSSEDECHAFKILKIFEVGFIQPGSVKYSTYNNDEKCTKRYHPDKGIGLLNKICIGNV 1489  
 QY 1396 CRPCEDGASGSHR-----HSSVIFIFCFKLLYME 1425  
 Db 1490 CRGAGETCSLNQERIDVPLQIEKACETNVYVYKTLRIIE 1533

RESULT 8  
 US-09-017-947-2  
 ; Sequence 2, Application US/09017947  
 ; Patent No. 6303754  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VOGEL, CARL-WILHELM  
 ; APPLICANT: BREDEHORST, REINHORST  
 ; APPLICANT: KOCK, MICHAEL  
 ; APPLICANT: FRITZINGER, DAVID  
 ; TITLE OF INVENTION: RECOMBINANT PROCVF  
 ; NUMBER OF SEQUENCES: 39  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
 ; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY  
 ; CITY: ARLINGTON  
 ; STATE: VA  
 ; COUNTRY: USA  
 ; ZIP: 22202  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent in Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/017,947  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/662,227  
 ; FILING DATE: 14-JUN-1996  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: OBLON, NORMAN F.  
 ; REGISTRATION NUMBER: 24,618  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 703-413-3000  
 ; TELEFAX: 703-413-2220  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1642 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-09-017-947-2

Query Match 12.2%; Score 894; DB 4; Length 1642;  
 Best Local Similarity 22.5%; Pred. No. 4,4e-65;  
 Matches 361; Conservative 295; Mismatches 672; Indels 276; Gaps 56;  
 QY 18 AALAV-----APGRPLVAPGILIRGAVITIGVELL-BHCSQVTVKA-----ELLKTKAS 67  
 Db 10 AALLIGFGSSHGALYTLITPAVLRTDTEQILVEAHGDSPTKQDILFVHDPKQKTLF 69

QY 68 NLTVSLAEVGEFEXKSFYTLTLDELPLNSADEIYELRVGTODEILFSNSTRLSEFTK 127  
 Db 70 QTRVDMNAGAMLVLT---PTIEIPAKEVSTSRQNVVVOVGTGQVLEKVLSSYQSS 126  
 QY 128 RIVSVIQTQDKLKYKQKQVFRIVTLFSDPFYKTSNLILKDKPSNLIQOMLSQOSDLG 187  
 Db 127 --FLFIQTQDKGIYTPGSPVLVRFVSMDSNTSRMNTVIVETQTEGILV--SNSVILN 181  
 QY 168 VISKTFQLSHPLIGDMSI--QOVNODQTYQSPOVSEYVLPKFEVTLQTP---LYCSMN 242  
 Db 162 FP-WPYNLPLDVSIGTMTIVAKIEHSPENNTAIPVRYKTVLPSEVRLQPSKEFYIDGN 240  
 QY 243 SKHLNGITTAKYTGKPKYKGVDTLTFPLSPWKKKNI---TKTFKINGSANFSPNDE 298  
 Db 241 -ENFHVSIYARVLAGEVEG--VAFVLFVGKIDDAKSIPLSLTRPLIDGOKATLKBDT 298  
 QY 299 MKNVMDSSNGISEYLDSSPGPVELITVTSVTSIGISNASTVNFKKQHDYIIEFDYTT 358  
 Db 299 FRSRPPNINELVGHITLAS-----VTVMTESGSDMVTTEQGIHIVASPYOHFTKTPK 352  
 QY 359 VLKSLNFTATVKTTRADGNQTLLEBRNNVITVTOENATYEWSSGNSGQKMEAVQKI 418  
 Db 353 YPKRMPVELTVYTNPDSPA-----AHVVSSEAF-----HSMGTLLSDGTAKL 398  
 QY 419 NYTVPOSGTFKIEPILEDSELOLKAYFLGSKSMVAHSLFKSPSKTYIQLKTRDENIK 478  
 Db 399 ILNTPLNQ---SLPITVTRNHGDLPRERQATKSMTALAYTOGSGSGVHLVAITSTBIK 455  
 QY 479 VGS--PEFLVYSGN---KRLKELSYNVVSGQLVANG-----QNSTMFSLPENSWT 526  
 Db 456 PGDNLPVNFVNVKGNANSLKQIKFTYTLINKKIFKQGRQRRDQGNLVYNNMLHTTDPDI 515  
 QY 527 PKACIVYIYIDDEGISDVAKIPVQ-----LVFK--NKIKLWVKYKAPSEVSVIRI 578  
 Db 516 PSFRFVAYIQVONNEIYADSVWVDKOTCMGTGLVVKGNLIQM-----PGAAMKTKL 567  
 QY 579 SVTOPDSIVGIYAVDKSVNLNANASNDITMENVVELBLVNTGYLYGMFNSFAVFOECGL 638  
 Db 568 E-GDGRAGVGLVADKAVYVNDKTKISQAKIMPTIEBDSGCTAGSQQNNLGVFEDAGL 626  
 QY 639 VWLTDANI--TYD----- 649  
 Db 627 ALTTSTNINIKQSAKACQOPANRRSSVLLDSNASKAEPQDQDLKCCEDVMBENP 686  
 QY 650 -----YIDGVN-MAEYARFM--BENE-GHIVDIHDF 678  
 Db 667 MGYTCERAKYIQRGADKAAFLCCRYIKGVNDENORESELFLARDNEDGFLADSDII 746  
 QY 679 SLGSSPHVKHFPETWIL-----DTNMGYRIYQEEFVTVPDSITSVAVTGFISEDLG 732  
 Db 747 S-----RSDFPKSWMLTDLTLEBRNSQGISKTSFYLRDSITTVVLAVSFTPTKG 799  
 QY 733 LGLTTTPVELQAFQPFIFLNLPSYVIRGEEFALEITIFNYLKATEKVIIEKSDKEDI 792  
 Db 800 I-CVAEPEYIRVMKVFIDLOMPYSVVNQEVEIRAILAHNVNEDIYVRVELVNPAP-- 856  
 QY 793 LMTSEINATGHOQTLVPSSEDGATVLPRIPTHG--EITITVTALESTPSTDAVTOML 850  
 Db 857 --CSASTGQRYRQFPFKALSSRAVPVIVPLBOGLHDEVIKASVDEALMSDGRKKLK 914  
 QY 851 VKAEGIEKYSQSLDLDTNRLQST-----LKTLSFPFPNVTGSEERQVITAIGVIL 904  
 Db 915 VPEGVQVQSIYITVYKLDPRAGVGTQLEVIKAKKLDKRDVDTETETKIIIIQSPVAOI 974  
 QY 905 GPSING--LASLIMPYGCGQNMNINPANI---YIIDLTKKKQLDTNLKELASPMQ 959  
 Db 975 ENSIDGSLANMLITTPSGCGQNMIRMAAPVIAITYVLDTEQMETLGINRTEAVNOIVT 1034  
 QY 960 GYQRELLYQREDGFSAPAGNVPDSGSTMASFVRCF--LEADPIYIDQDQVLAHTTYML- 1017  
 Db 1035 GYAQQMVYKADHSHYAAFTN--RASSSWLTAYVVKVAPMAKMAVAGISHETIIICGVRLI 1092  
 QY 1018 KGHQKSNGEFMDPGRVHISELOG--GNKSPVTLTAVITVSLGVR-----KYQGNID--V 1068





QY 851 VKAEGIEKYSOSILDLTDNRLOST-----LKTLSFPPPTVTGSRVOITAIQDVL 904  
 Db 915 VVEBQKSVITVIVKIDPRKAGVGGTQLEIVIKARKLDDRVPTLETKIIIOGDPAQI 974  
 QY 905 GPSING--LASLIRMPYGGGEONMNFAPNI---YILDVLTKKQUTDLKKKALSFMQ 959  
 Db 975 ENSIDGSKLNHLITTPSGGGEONMTRMAPVATVYLLDTTEQWETLGINRRTKTEAVAQI 1034  
 QY 960 GYORELLYQREDGSPFAGNNDPSGSTMVSAFVLRCF--LEADPYIDIDQNVHRTYTWL 1017  
 Db 1035 GYAQQMVYKKADHSYAAFTN--RASSWLTAVVYKVFAMAAKAVAGISHEIITCGVRW 1092  
 QY 1018 KKHQSNHGFMDPGRYVISELOG--GNKSPVLTAVYVTSILGR---KYQPNID--V 1068  
 Db 1093 INRQOPDGAFFKFNAPVLSGTMOGIGAGAEFVYLTAFILVALLSEKTIQNDYVNSIDSSI 1152  
 QY 1069 OESIHFLSEFSEFSGISDNVTLALITVYALSVSQSPRAKELNMLTRABEGGMOFVSS 1128  
 Db 1153 KKAIVYLLKKYRK--IQRPYTTALTAVALA-----AADLNDRVLMASGRDHV--- 1201  
 QY 1129 SKLSWOPRSIDIEVAAYALSHFLOFOTSEGIPIRWLSRQNSLGGFASGTODPTV 1188  
 Db 1202 ---BEVNAHTNIEGTSYALALMKKKFDOTGPIVWLTQNTYGETYQOTATWAF 1257  
 QY 1189 KALSEFPAALMTER--TNIQVTGSPSPPL-----AVQP---NAVVISANG 1232  
 Db 1258 QALAEVIEIOMPHTKDLNLTITELPDRBVRIRYRINVENALLARVETELNDIYVTSAG 1317  
 QY 1233 FGFAICQNVVYVNVKASGSSRRRSIQNGEAPDLVAVKENDLNHD-----LVNCT 1286  
 Db 1318 DOKAMTITTFYNAOL---QEKAVCNK--PHLVNVE--NHNANMAGKALMLKICT 1369  
 QY 1287 SFGSGRSGMALMEVNLISGFVWSEALS-----LSEVYKVEYDHGKLN-----LYLDS 1336  
 Db 1370 RYLGVDSTMTIIDISMLTGFDPDAEDLRLSKVDRIYKREYVDNNAQKAVIITLAK 1429  
 QY 1337 VNEIQ--FCVNIIPAVENFKVSNTOASVSIIDYEPBRQAVRSYSEVYKLSCDLSPVQ 1395  
 Db 1430 VSHSEDECHFKILKHFEVGFIOGSAVKYSYNIDCKTKFYHDKGTGLINKICIGNV 1489  
 QY 1396 CRPCDDGASGSH-----HSSVFICFCFLYME 1425  
 Db 1490 CRACGETCSLNHQRIDVPLQIEKACETMNDVYVYKTLRIE 1533

RESULT 10  
 US-08-662-227-35  
 ; Sequence 35, Application US/08662227  
 ; Patent No. 5922320  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VOGEL, CARL-WILHELM  
 ; APPLICANT: BREDEHORST, REINHORST  
 ; APPLICANT: KOCK, MICHAEL  
 ; APPLICANT: FRITZINGER, DAVID  
 ; TITLE OF INVENTION: RECOMBINANT PROCVF  
 ; NUMBER OF SEQUENCES: 39  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
 ; ADDRESSEE: P.C.  
 ; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY  
 ; CITY: ARLINGTON  
 ; STATE: VA  
 ; COUNTRY: USA  
 ; ZIP: 22202  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/662,227  
 ; FILING DATE: 14-JUN-1996  
 ; CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:  
 ; NAME: OBLON, NORMAN F.  
 ; REGISTRATION NUMBER: 24,618  
 ; REFERENCE/DOCKET NUMBER: 1126-0107-0X  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 703-413-3000  
 ; TELEFAX: 703-413-2220  
 ; INFORMATION FOR SEQ ID NO: 35:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1648 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-662-227-35

Query Match 12.0%; Score 880; DB 2; Length 1648;  
 Best Local Similarity 22.7%; Pred. No. 6,6e-64;  
 Matches 365; Conservative 295; Mismatches 664; Indels 284; Gaps 59;

QY 18 AALAV-----APGRFLVTAQIIRPGNVITIGELL--EHCPQVTVKA-----ELKLTAS 67  
 Db 10 AALLIGPSSSHGALYTLITPAVAKTDEQILVEAHGDSPPKQDLIFVHDFPRKQKTLF 69  
 QY 68 NLTVSVLBAEGVFEKGSFKTLTSLPLNSADEIYELVTRTORDEILFSNSTRLSFETK 127  
 Db 70 QTRVDMNAGGLVY---PTIEIPAKVSTSRQNYVVGVTGQVLEKVLISYQSS 126  
 QY 128 RISVFIQTDKALYKPKQEVKFRIVTLFSDPKRYKTSNLIL--KDPKSNLIQOWLSQSD 185  
 Db 127 --FLFIQTDKGIYTGSPVLRYVFSM--DHHTSKMKNTVIVIEFQPEGILVS---SNSVD 179  
 QY 186 LGVISTKQLSHPIITLPGMSI--QYVNDQYVYQSFQSEVYLKPEFTYLTQF---LYCS 240  
 Db 180 LNFF--WPNLDPVLSLGTWRIVAKYEHSPENVYAVDRKXVLLPSEFVRLDQSEKPFYTD 238  
 QY 241 MNSKHLNGTITAKYVYGRVGDVTLTFLPLSFMGKKNI---TKTPKINGSANFSND 296  
 Db 239 GN--ENFHSITARYIYGEVEG--VAFVLFYKIDAKSIPDSLTRITIIDGKATLYR 296  
 QY 297 EEMKNVMSNGLSLEYDLSSPGVYELLITVYESVTGISRVNSTVNFKQHDYIIEFDY 356  
 Db 297 DTFRSRFPNLNBLVCHTLVYAS-----VTWMTSGSDMVWTEQSGIHIVASPYQIHFTKT 350  
 QY 357 TTVLAKPSLNTFYATYKVRADGNQLTLEBRNNVITVYQRYTERVSSGNSGNGQMEAVQ 416  
 Db 351 KYEKPQMPYELVYVYVNPDSPA-----AHVPVSEAF-----HSMGTLISDGT 396  
 QY 417 KINYVPOSGTFKIEFPILEDSELOLKAYFLGSKSSMAVSHLFSKPSKTYIQKTRDEN 476  
 Db 397 KLILNIPINAQ---SLPTVYRTHGDLPRERQATKSMTAIAVYQGGSGNYLVAITSTE 453  
 QY 477 IKVGS--PELVVSGN-----KRLKELSYMVVSRQLVAVGKONS-----TM--FSLTPE 522  
 Db 454 IKPGDNLVFNENVVNGNANSLSQIKYFTYLLINKKIKFVNGQPRRDGNQVLTMLNHTIPD 513  
 QY 523 NSWTPKACVITYIYEDGEIISDVAKIPVQ-----LVFK--NKIKLWSYKAKPSEK 574  
 Db 514 LIPSPR--FVAYYQGNNEIIVADSVWVDKOTCMGTILVVKGDNLIQMGAAKIKILEGDF 571  
 QY 575 SLRISVTPDSIVGVIVADKSVNLMNASNDITMVNVHLELYNTGYLLGFMNMSFAVQ 634  
 Db 572 GAR-----VGLVADKAVYVYVNDKIKISQAKIMDTLEKSDPGTAGSGGNLGVFE 622  
 QY 635 ECGLMVLTDANL--TKD----- 649  
 Db 623 DAGLALTSTYLNKQSSAAKCPQANRRRSSVLLDSNASKAAEPQDQLRKCCEDVM 682  
 QY 650 -----YIDGYVD--NAFYARFV--EENE--GHVVD 674  
 Db 683 HENPMGYTCERAKRYIOGDACKAFLCCRYIYGVDENOREBELFLARDNDNDEGFAD 742  
 QY 675 IHDPSLSSPHVRKHFPETWIL-----DTNMGYRIYQIEFVYVVPDSITSVATGFVIS 728

```

Db 743 SDIIS-----NSDFKMMMLTKDLTEEPNSGJSSKTSFYLRSJTTTWWLAVSFT 795
Qy 729 EDLGLGTTTPEVLOAFOPFFIFPLNLPYSYIRGEFALETITFNLYLKADATEVKIIEKSD 788
Db 796 PTGDI-CVAEYELRWKVFIDLOMWYSYVKNQVIRAILNHYVEDIYVRELLYNP 854
Qy 789 KFDILMTSSSEINATGHOQTLVSEDAVYLPFRPHLG--ELPTVTALSPASDAVT 846
Db 855 AF-----CSASTKGORRQOPPIKALSSRAVFPVLPLEOGLHDEIKASVQOEALMSDVR 910
Qy 847 OMILVKAEGIEKSSOSIILDLTONRLQST-----LKTLSFSPPTVYGSERVQTAI 900
Db 911 KKLKVPBEGVQKSLVTIVKLDPRKAVGGTQLEVIKARKDDRRVDTETIKTIIIOQDPV 970
Qy 901 GDVLGSPING--LASLIRMPYGCCEQOMINFAPNI--YILDYLTKKKQOTDLKELKALS 955
Db 971 AQIENIDSKLNHLITPSGCEQOMIMAPVIAVITYIDTTEQMETIGINRTEAVN 1030
Qy 956 FMRQYORELLYOREDSFSAFNGYDPSGTWLSAFVLRCP--LEADPYIDIDQNVLRHY 1014
Db 1031 QIVYGAQOMVYKKADHSYAAFTN--RASSMTLAAVYVYFAMAKKAVAGISHIICGV 1088
Qy 1015 TWL-KGHOKNGEPMWDCGRVYHSELQ--GNSKPYLTAYIVYSLGVR---KYQPN 1066
Db 1089 RMLILNRQOPDGAKEKAPVLSGTWQGGIOGAEVEVYLTAFIVALLSKTICNDYVNSL 1148
Qy 1067 D--VQBSIHFLSEFSRGISDNVYLTALITYALSVSQPKKEALNMLTWBAEOEGMQPW 1124
Db 1149 DSGIKKATNVLKRYEK-LQRPYTTALAYALA-----AADQINDRVYLAASSTGDDHW 1201
Qy 1125 VSSEKSKDSMOPRSLDIEVAVALSHLQFQTSSEGIPIRMWLSRGRNSLGGFASQDT 1184
Db 1202 -----EENATNHNIEGYSYALLMLKKMKFDQGPVIRWMLTDQNFGETYGOQAT 1253
Qy 1185 TVALKALSEFPAALMNTER-TNIOVTVYGPSSPSPL-----AVQOP---NAVNI 1228
Db 1254 VMAFOALAEYELQPHKDLNLDITIELPREVPIRIRYENALARTVETKLNODITV 1313
Qy 1229 SANGFGFALCOLNVVYVKAAGSSRRRSIONOEAFDLDAVAKENKODLNHD-----L 1282
Db 1314 TASGDGKATMTILTFYNAQL-----QEKANVCNK--FILNVSVE--NHLNMAKAKGALML 1365
Qy 1283 NCTSFSGPRSGMALMEVNLGFWVPSEAS-----LSEYVKEVYDHGKLN-----L 1332
Db 1366 KICTRYLGEVDSWTITIDISMLTGLFDADLTRLKSGVDRYISRYEVDNMAQKAVAIL 1425
Qy 1333 YLDSVNETQ-FCVNTIPAVRNFKVSNTODASYIVDYEPFRQAVRSYNSSEVKLSSCDLCS 1391
Db 1426 YLNKVSHEDECLHPKILKHFEGVFIQGSYKVSYNLDEKCTKFPHPDKGTGLANKIC 1485
Qy 1392 DVQGCRCFCEQDQASGSH-----HSGVIFFCFKLLYFWE 1425
Db 1486 IGAVCRCAGETCSLNHQERIDVPLQIEKACETNVDVYTKTLRIEE 1533

```

## RESULT 11

```

US-09-017-947-35
; Sequence 35. Application US/09017947
; Patent No. 630354
; GENERAL INFORMATION:
; APPLICANT: VOGEL, CARL-WILHELM
; APPLICANT: BREDEHORST, REINHORST
; APPLICANT: KOCK, MICHAEL
; APPLICANT: FRITZINGER, DAVID
; TITLE OF INVENTION: RECOMBINANT PROCV
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA

```

```

; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Releasee #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/017,947
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/662,227
; FILING DATE: 14-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 1126-0107-0X
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1648 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-017-947-35

```

```

Query Match 12.0%; Score 880; DB 4; Length 1648;
Best Local Similarity 22.7%; Pred. No. 6,6e-64;
Matches 365; Conservative 295; Mismatches 664; Indels 284; Gaps 59;

```

```

Qy 18 ALAY-----APGRRLVAPAGIIRGCVNTIGVELL-EHCSPQVYVKA-----ELKTKAS 67
Db 10 AALLIGPSSSHALTYLLTPAVLRQTEBQILVEAHGSTPKQDLTFVHDFPRKQKTLF 69
Qy 68 NLTVSVLEAGVEFEKSGFKTLTLPSPLSADBIYELRYGRTODELFPNSSTRLSPEK 127
Db 70 QTRVDNMPGKMLYT---PTIEIPAKEVSTDSRQNVYVVOYTGPOVRLKEKVLSSYQSS 126
Qy 128 RISVFQTDKALYKPKQEVKFRIVTLFSDFKPKYKTSNLILI--KDKKSLIQQWLSQQSD 185
Db 127 --FLFIQTDKGYITPSSPVLYRVFSM--DHHTSKMKTYIVERTQTEGLIVS---SNSVD 179
Qy 186 LGVISTPQLSSPILIGDMSI--QVQVNDQTYQSQVSEYVLPKREYVLTQTP---LYGS 240
Db 180 LNEF-WPYNLPDLVSLGTWRIVAKYHSPENYAYVDFVRKYVLPSEVRLQSEKEFFYID 238
Qy 241 MNSKHLNGTITAKYTYGKPEVKGVDVTLPLPLSPWGGKKNI-----TKTFKINGSANFSND 296
Db 239 GN-ENFHVSTITARYLYGEEVGG-VAFVVLGVKIDAKKSIPLSLRIPILIDDGKATLKR 296
Qy 297 EEMKNVDSNGISEYLDSSPGVPEILTTVESYVIGSRNVSTNVFPKOHDIIEFPDY 356
Db 297 DFRSRFPMLNELVGHLYVAS-----VTVMTESSGDVVTTEQSGIHIVASPOUHFYTK 350
Qy 357 TTVLKSPLNFATYKTRADGNOLTEERRNVVITYORANTEYVSGNSGNGQKEAVQ 416
Db 351 PKYFKGMEYELTVYTNPDGSPA-----AHVAVSEAF-----HSMGTLLSDGTA 396
Qy 417 KINYVPOSGTFKIEPILEDSELOLKAYFLGSKSKSMVHSLFKSPKTYIOLKTRDEN 476
Db 397 KLILNIPLNAQ--SLPTVYRNHGLPREQATKSMTALAIVOTOGSGNYLHVALTSTE 453
Qy 477 IKVGS--PEELVSGN-----KRLKELSYNVVSRGQVAVGKONS-----TN-FSLTPE 522
Db 454 IKRGDNLPVNFVNVKGNANSLKQIKYFTYILNKGIFKYGORPRDQOMLVYTNLHITPD 513
Qy 523 NSMTPRACIVYVYIEDGIIISVLKIPVQ-----LVFK--NKILYNSKYVAESESKEV 574
Db 514 LIPSPR--FVAYYQVGNNEIVADSVWVDVDTCMGTLLVKGDLIQMPGAMKIKLEGDF 571

```

575 SLRISTQDSIVGIYAVDKSVNLMASNDITMENVHELELYNTGYI.GMFNNSFAVFO 634  
 572 GAR-----VGLVADVKAAYVLDKRYKISQAKIMDTIEKSDFCGTAQSGQNNIGVE 622  
 635 ECGLWLTJLANL-TTQ----- 649  
 623 DGLALTTSTNLNTKORSAKCPQANRRRRSSVLLDSNASKAAEFQODLRKCEDVM 682  
 650 -----YIDGYD--NAEYAEERFM--EENE-GHIYD 674  
 683 HENPMGYTCERAKXYIQEGBDACKAALFECRCYIKGRDENQRESEFLARDNEDGFIAD 742  
 675 IHDPSLGGSPHVRKHPETWIML-----DTNMGYRIYQEFVTVYDSTISWATGCVIS 728  
 743 SDIIS-----RSDFPKMWMLTKDTEBPNSQGISSTMSFYLADSTITWVVLAVSPT 795  
 729 EDGLGLTTTPVELAOFQFPFITFLN.PYVIRGEFALETITFNILKDAITEVKVITIEKD 788  
 796 PTKGI-CVAPEYRIRVMKVFIDLOMPYSVVKNQVEIRAIILHNYNEDIYVAVELLNYP 854  
 789 KFDIMTSEINATGHQOTLVPSBEDGATVLPPIRPTHIG--EIPITVALSPTASDAVT 846  
 855 AF-----CSASTGQGRKQOPPIKALSSRAVPFVYVLEQGLHDVEIKASVQBALMSDGV 910  
 847 QMILVKAEGIEKYSQSILLDLTDNRLOST-----LKTLSFSPFNVTGSERVQITAI 900  
 911 KKLKTVBEGVQSIYIVTKLDPRAKGVGTQLEVIKARLDRVPDTEIETKIIIGDVP 970  
 901 GVLGSPSING--LASLIRMPYCGEONMNFAPNI--YILDYLRKKKOLTNLKEKALS 955  
 971 AQIIEISIDGSKLNLHILITPSGGEONMIRMAFVATVYLLDTTEEMETLGINRTEAVN 1030  
 956 FROGQYREBLKYOREBDSFSAFCNVPDPSGSTMASAFVLCF--LEADPYIDIDQNVLRHY 1014  
 1031 QIVTGAQOQVYTKADHSTAAFLTN--RASSSWLTATVAVVAFMAAMVAGISHEIICGV 1088  
 1015 TWL-KGHOKSGNEFMDPGRIYHSELQGG--GNKSPVTLTAVYVTSILGR-----KYQPI 1066  
 1089 RWLILNRQOPDGAFKENAVPLVSGTQGGIQAABEYVLTAFILVALLBESKTCINDVYNSL 1148  
 1067 D--VOESIHFLSESESRGISDNTTALITVALSSVGSPPAKALANLTVRAOEGMGQWF 1124  
 1149 DSISIKATNYLLKTYEK-LQRPYTTALTAYALA-----AADQINDRVILMAASTGRDH 1201  
 1125 VSESELSQSPRSLDIEVAAYALLSHLOFQTSRGITIMWLSQGRSLGFASTQDT 1184  
 1202 -----EYNAHTNIEGTSTALLALMKKFDQGTGPIVWMLTDQNFYGETYQTOAT 1253  
 1185 TVALKALSEFALIMTER-TNIQVTVYGSPPPL-----AVVOP--MAVNI 1228  
 1254 VWAFOALAYEIQMFPHKDLNDITIELDREVPRIYRINERALLARVEYRKLANDITV 1313  
 1229 SANGFQALCOLNVYVNVKASGSSRRRSIQOEAFDLDAVAKENDDLNHYD-----L 1282  
 1314 TASGQKATMTILTFYNAQL--QEKANVCNK--PHLNVSE--NHLNANGAKGALML 1365  
 1283 NVCTSFSGRGSGMLMEVNLISGFVVPSEALS-----LSEYVKKYVYHGLKN-----L 1332  
 1366 KCTSRLEGVDSTMTIIDISMLTGLPDAEDLTRLSKGDYRISRREVENNNAOKVAVIT 1425  
 1333 YLDSVNETO--FCVNIPAVNFKVSNTOQASVIVDYBPRQAVRSYNSGVKSLSDLS 1391  
 1426 YLANKSHSDECLHFKILKHFEVGFQPSVKVYSTYINLDEKCTKYHHDGKTGLINKLC 1485  
 1392 DVQGRCPCEDGASGSH-----HSSVIFICFKLLYFME 1425  
 1486 IGVNCGAGETCSLNHGERIDVPLQIEKACETNVYVYKTKLIRREE 1533

RESULT 12  
 US-08-447-411-76  
 ; Sequence 76, Application US/08447411

Patent No. 5773243  
 GENERAL INFORMATION:  
 APPLICANT: FRITZINGER, DAVID C.  
 APPLICANT: BREDEHORST, REINHARD  
 APPLICANT: VOGEL, CARL-MILHELM  
 TITLE OF INVENTION: DNA ENCODING COBRA C3, CVF1, AND CVF2  
 NUMBER OF SEQUENCES: 81  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: P.C.  
 STREET: 1755 S. Jefferson Davis Highway, Suite 400  
 CITY: Arlington  
 STATE: Virginia  
 COUNTRY: U.S.A.  
 ZIP: 22202  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent in Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/447,411  
 FILING DATE:  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/043,747  
 FILING DATE: 07-APR-1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: OBLON, NO. 5773243man F.  
 REGISTRATION NUMBER: 24,618  
 REFERENCE/DOCKET NUMBER: 1126-101-0  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (703) 413-3000  
 TELEFAX: (703) 413-2220  
 TELEX: 248855 OPAT UR  
 INFORMATION FOR SEQ ID NO: 76:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1333 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-447-411-76  
 Query Match 10.5%; Score 768.5; DB 1; Length 1333;  
 Best Local Similarity 24.4%; Pred. No. 9.7e-55;  
 Matches 301; Conservative 208; Mismatches 477; Indels 247; Gaps 45;  
 349 YIIREFDYTVLKPSELNFTATVKTTRADGNQVTLERNNVITVTOHNYTEYWSGNSG 408  
 22 YQIVTTPKPKFKPCMPPELYTVYTKPDGSPAA-----HVPVSEAIHSE-----G 67  
 409 NOKMEAVQKINTVYQSGFTKIEPFILEDSELOKAYFLGSKSMAVSLFKSPSKTYI 468  
 68 TLLSGTKLFLNTPQNAQ--SLPITVRYHGDIPREROAIKSWTATAYOTQGGSGNYL 124  
 469 OLKTRDENIKVGS--PFLVLYSGN--KRLKELSYMYVVSQGLVAVGKONS-----T 515  
 125 HVAITSTIKREDNLVAVNVRGNANSJLNQIKYFTYLLINKGKIFKVRQHRGSGENILVT 184  
 516 M-FSLTPENSWTPKACVIVVYIEDDEITSDVLKIPVQ-----LVFK-----NKIRLY 562  
 185 NMLHITPD--LIPSRFAVYQVGNELVADSVWVDVDTGCTGLVYVAGATSRDRIGM-- 241  
 563 WSKVKAPESEKYSRLISTQDSIVGIYAVDKSVNLMASNDITMENVHELELYNTGY 622  
 242 -----FGAAMKITLE-GDPGAWIGLVAADVKAAYVLDKRYKISQAKIMDTIEKSDFC 293  
 623 LGMFNNSFAVFOEGSLWLTJLANL----- 647  
 294 AGSGQNNIGVFEDDGLALTTSTNLTKORSAKCPQANRRRRSSVLLDSNASKAAQFO 353  
 648 -----KDYIDGYDND--AEYAEERFMEENEG-----HIV----- 673

```

Db 354 DDLRKCCEDEMGHENPMGHTCEKREKXIOEGDACKAFLCECHYIKGIQDNKRESEFL 413
Qy 674 ---DIHPSLGSSPHV-RKHPEETIWL-----DTNMGVRIQEFVTVPDSITSM 720
Db 414 ARSDFEDDLFEGENITSRDSEFESWLMEOLESHPNKGISKIVPY---LDSITTW 470
Qy 721 VATGFVISEDGLGITT-----PVELQAPQPFIFLNLPSYVIRGEFALEITIFNY 773
Db 471 -----ELLAVGLSTPKICVAPEYETVMDFIDLOLPYSVVKOEQKIRAVLVNY 522
Qy 774 LKDATEVVIIEKSKDFILMTSSFINATGHQOQLLVPSBDGATVLPFIRPHLG--EIP 831
Db 523 ADKOIYVVELLYSPAFCSASTESQ---RYRQOLPIKALSRRAVSFVIVPLEGGLHVE 578
Qy 832 ITVTALSPASDAVOTMLVKAEGSEKYSOSILLDLTDNLQST-----LKTLSFSGP 885
Db 579 VTAASGGLMDGVKKIKVYPEGEMKSIIVTIELDPHTKIGIGQVAVLANKLNDRV 638
Qy 886 PNTVTSERRVOITAGDVLGSPSING--LASHIRMEPYGGEQNMINF-APNT--YILDYLT 940
Db 639 DTEIETKITIIGDPAVQTIENSIDGSKLNLIFPFEGGEQNMIRMTAPVATVYLDITQ 698
Qy 941 KKKQUTNLKKALESFMAQIGQRELLVQREDSBSAFGNVPSGSTWLSAVLACF--LEA 999
Db 699 QMETIGIRRTAVVQIMTGYAQQLVYKKAHSYAFTN--SASSWLTAVVXIFALAA 756
Qy 1000 DPYIDIDONVLRHTYTWL-KGHOKSNGBFMDPGVRIHSELQGNKSPV--TLRAYITS 1055
Db 757 KIYKDINHEIYCGGKRWILNRQTDGVRENAPVLFCTWGGGIGQABESGLTAFILVA 816
Qy 1056 LLGYRKQPNIDVOESIHLESEFSG-----ISDNYTLALITVALSVGSPKAK 1105
Db 817 LLESRSI-----CNAYINILDSISISKANDVYLKKYEKQIRYTTALTAYALAA-----AA 865
Qy 1106 EALNMLTRARQEGGMOQWVWSESEKLSQPSRIDIVAAVYALISHLOQTSBGLPIM 1155
Db 866 ERLNDRVLMAASTGRNMEBPNA-----HTNIBGTSYALALAKMKKPEAGPV 917
Qy 1166 RWLSRORSLSGFASTODTPTALAKALSFALNMTER-TNFOVTWTPGSSPSP----- 1217
Db 918 QMLIDQYGGYGGYGTOTATVMMFOALAYEIQMFKHKLNDITIELPREVPRIRYRINY 977
Qy 1218 ---LAVVQPMVAVN---ISANGFGFALCOLNVVYVNVASGSSRRRSIQOQAFDLDVA 1269
Db 978 ENALLAQVETBKLMEDFTVASGQKATMTILTYNNAQL-----REDAVNCK--FHLDS 1031
Qy 1270 V-----KXNKDILNHYDLNVCISFSGPGRSGMALNEVNLISGMPWSEASIS--LSETVK 1321
Db 1032 VENVOJLNKEAKGAKGALKIKICTRYLGEVDSVTMTIIDVSMLTGFVPDTEJLRSKQVD 1091
Qy 1322 K-----VEYDH-----GKLNLYLDSVNETQ--PCVNI PAVRNPKVSNTOASVSIYDYER 1371
Db 1092 RYISMFELNNMAQKGIYIITLYDKVSHSEDECLHFKILKHEVGHIOGSKVVSYNLD 1151
Qy 1372 RQAVRSYNSVYKLSGCDLCSDOVGRPCEDGAS 1404
Db 1152 ECKTKIYHPDEATGLNKICVGNVCRCAEFICS 1164

```

```

RESULT 13
US-08-662-227-34
Sequence 34, Application US/08662227
Patent No. 5923230
GENERAL INFORMATION:
APPLICANT: VOGEL, CARL-WILHELM
APPLICANT: BREDEHORST, REINHORST
APPLICANT: KOCK, MICHAEL
APPLICANT: FRITZINGER, DAVID
TITLE OF INVENTION: RECOMBINANT PROCVF
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESS: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
ADDRESS: P. C.

```

```

STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/662,227
FILING DATE: 14-JUN-1996
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 1126-0107-0X
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 1333 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-662-227-34

Query Match 10.3%; Score 758.5; DB 2; Length 1333;
Best Local Similarity 24.4%; Pred. No. 6,6e-54;
Matches 301; Conservative 208; Mismatches 477; Indels 247; Gaps 46;

Qy 349 YIIEFFPYTVLKLPSLNFITVYKTRADGNQLILERRNNVYTVQRNTEWSSGNSG 408
Db 22 YQIFFTPKPKFKMGMYELTYVTKPDGSPA-----HVPVSAIHSE-----G 67
Qy 409 NQKKEAVOKINVTYVQSGTFKIEFPILEDSELOLKAYFLGSKSMVHSLFKSPSTYI 468
Db 68 TTLSGDKAKLFLNTPQNAQ---SLPITVRKNHGDLPREPAIISMRTAVATQGGSGNYL 124
Qy 469 QLKTRDENIKVGS--PPELVNQN---KRLKELSYNVVSRGOLVAVGKNS-----TM 516
Db 125 HVAITSTEIKGDNLPVNFVNRGANSLNQIKFTYILINKKGIKFKVGRQHRGDNLVTM 184
Qy 517 -FSLTPENSMTPKACVYIYIEDGEL-ISDVLKIPVQ-----LVFK-----NKIKLY 562
Db 185 NLHITPD--LIPSRFVAYVQVGNNELEVADSVWVDKOTCMGTLVVKGATSRDNRQM- 241
Qy 563 WSKYKAPSEKVSRLISVTQPSIVGIVAVDKSVNLKMASNDITMEVVAHELEYNTGY 622
Db 242 -----PGAMKIKLE-GDPGAMIGLVAVDKAEVYLNDRKISQAKIMWTIEKSDGCT 293
Qy 623 LGEMFNFAVFORBGIMVLTDANLT----- 647
Db 294 AGSGQNNLGVFEEDAGLALTTSTNLNTRKQBSAKCPQPNRRRRSSVLLDSNASKAAQFQ 353
Qy 648 ---KQYIDGVYDN---AEYARPMENEG-----HIV----- 673
Db 354 DDLRKCCEDEMGHENPMGHTCEKREKXIOEGDACKAFLCECHYIKGIQDNKRESEFL 413
Qy 674 ---DIHPSLGSSPHV-RKHPEETIWL-----DTNMGVRIQEFVTVPDSITSM 720
Db 414 ARSDFEDDLFEGENITSRDSEFESWLMEOLESHPNKGISKIVPY---LDSITTW 470
Qy 721 VATGFVISEDGLGITT-----PVELQAPQPFIFLNLPSYVIRGEFALEITIFNY 773
Db 471 -----ELLAVGLSTPKICVAPEYETVMDFIDLOLPYSVVKOEQKIRAVLVNY 522
Qy 774 LKDATEVVIIEKSKDFILMTSSFINATGHQOQLLVPSBDGATVLPFIRPHLG--EIP 831
Db 523 ADKOIYVVELLYSPAFCSASTESQ---RYRQOLPIKALSRRAVSFVIVPLEGGLHVE 578

```



```

Dh 699 QWETLIGIRRRREAAVNVQIMTYAQOVLVYKKAHSAFAFN--SASSSWLTAIVYKIFALLA 756
Qy 1000 DPYIDIDONVTLHRTYTWL-KGHQKSNGEFMDPRYVISELOGCNKSPV--TLTAIYTS 1055
Dh 757 KIYVDINHEIYCGGMRWMLINRQRFDGYFRENAPVLFGTMQGGIOGAPEEGSLTAFILVA 816
Qy 1056 LLGRKYQPNIDVQESIHFFLESEFRG-----SDNVTALITLYLSSVSPKAK 1105
Dh 817 LLEBSRT-----CNAYINILDSISKATDYLLIKTEKQRPPTLALTYALA-----AA 865
Qy 1106 EALNMLTWRAQEGGQMFWVSESKLSDSWOPRSDIEVAAYALSHFLQFOTSEGDIM 1165
Dh 866 ERLINDRVLMAASTGRNWEENPNA-----HTNINIGTSLALLALMKKKFVAGBPV 917
Qy 1166 RWLSQRKSLGCFASDTODTYALKALSEFALAMTER-TNIOYTYGSSSP-----1217
Dh 918 QMLIDQOYGGYGGTQATVMMFQMLAEYELQMPHKLINDITIELDREVPYRIRINY 977
Qy 1218 ----LAVQPAVNV-----ISANGPFAIQLNVVVNVKVASGSSRRRRSIOGEAPDLVA 1265
Dh 978 ENALLAQVETVEKLNEDFTVASGDDKATMTTITLVNAQI---RDAAVCKN--PHLVS 1031
Qy 1270 V-----KENKDINLHVLDNVCTSPSGPGRSGMALMEVNLISGPMVSEALS-LSETVK 1321
Dh 1032 VBNVQNLNKEKKGAGALKLKICTRFLYBVDSTWMLIIDVSLTGVVPPTEDLTRSKVD 1091
Qy 1332 K---VEYDH-----GKLNLYIDSVNETQ-FCVNIPAVNRKVSNTODASVSYDYEP 1371
Dh 1092 RYISWFEINNNMAOKGTVIYLDKXSHSEDECLHFKILKHFEPVGIQGSVYKVSYYMID 1151
Qy 1372 RQAVRSYNSVEYKSLSCDLCSVDQGRPCPEDGAS 1404
Dh 1152 ECKTIIHPDEATGLINKICVGNVCRABETCS 1184

```

RESULT 15-  
 US-07-779-172A-3  
 : Sequence 3, Application US/07779172A  
 : Patent No. 532838  
 : GENERAL INFORMATION:  
 : APPLICANT: Silberstein, David S.  
 : APPLICANT: Minkoff, Marjorie  
 : TITLE OF INVENTION: INHIB A Factor Which Inhibits  
 : TITLE OF INVENTION: Cytokine-Activated Leukocytes  
 : NUMBER OF SEQUENCES: 3  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
 : STREET: 1225 Connecticut Avenue, N.W.  
 : CITY: Washington  
 : STATE: D.C.  
 : COUNTRY: USA  
 : ZIP: 20036-2678  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: PatentIn Release #1.0, Version #1.25  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/07/779,172A  
 : FILING DATE: 19911016  
 : CLASSIFICATION: 514  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: Fox, Samuel L.  
 : REGISTRATION NUMBER: 30,353  
 : REFERENCE/DOCKET NUMBER: 0627.1270001  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: (202) 466-0800  
 : TELEFAX: (202) 833-8716  
 : TELEFAX: 248636 SSK  
 : SEQUENCE CHARACTERISTICS:  
 : SEQUENCE FOR SEQ ID NO.: 3:  
 : LENGTH: 645 amino acids

```

;      TYPE: AMINO ACID
;      TOPOLOGY: both
;      MOLECULE TYPE: peptide
US-07-779-172A-3

```

Query Match	Score	DB 1	Length
4.58	329	645	

Matches 143; Conservative 133; Mismatches 270; Indels 142; Gaps 22;

```

OY 26 PRFLVTLAAGILR-----PGGNVTIGVLELH-----CPQOVYAKAILKTA 66
Dh 2 PMSIITPNILRLSEETMVLAEADQODVPYTVIHFPGKULVLSSEKTVLTPATNMH 61
OY 67 SNLIVSV-LEAEVAFPEKGSFKTLLPLSP.LMSADEIYELRVTGRQ--DELLFSNSTRLS 123
Dh 62 GNVTFTIPANEFKSEKRNKPFVT-----VQATFBOVQVEKVL-----VS 102
OY 124 PETRISVFIOTDICALYKPOEVEFRIVTLFSDPEFKYTSINILIKDPSKNLLOQ-WLSQ 182
Dh 103 LQSGYL-FIOTDKITITPGSTVLXRIPTVNHKLLPBRVTMMNLENBEGIPVKDSSLSS 160
OY 183 QSDLVISKTFQLSHPILGDWSIQOVOND--QTYVOSFOVSEVYLPRFEVTLQ-TP.L 237
Dh 161 QNQGVLPLMDIDELVNMGMKIRAYEENSPQVSTFEVKEVYLSFEVIVPTEKF 220
OY 238 YCSNMSYHNTITIRAKTYGKPYKGVDTLTF-----LPLSFWKXKN 280
Dh 221 YYIYNEKGEVLTITARFLYGRKVGTAFAVIGIODBORISLPSLKRIPIEDSGEYVL 280
OY 281 TKTEKINGSANFSFENDEEMKNVMSNGLSEYLDLSPBPVEILTTVTESVYTSIRNVST 340
Dh 281 SRKVLVDQVQLRABDLVGKSLYSA-----TVLHSGSDMVQAERS 322
OY 341 NVFPKQNDYILIEFDYTVLKP.SINAFRTAVYTRADGN---QTLIEBRNNVYITVQRN 397
Dh 323 GIPVTSFYQHFKTPKPYFKPGMFDLMVTVTPDDGSPAYRVPAVQAGEDEVQVSLTOD 382
OY 398 YTEYMSGNSNGQKMEAVQINTYVPOSGCFKIEPIELDESSSELOLKAYFLGSKSMVNH 457
Dh 383 GVAKLSINTHSQK-----PLSITRTKKQELSEANQ-----ATRMQALP 423
OY 458 SLFSPSKTYIQLKTRDENIKVGSPPFELV-----SGNKELKELSNVWVSRGOLVANG 511
Dh 424 YSTVGNSSNNYHLBSVLTTTELPRGETLNVFLLRMDRAEAKIRYTVYIIMNGKRLKGR 483
OY 512 -----ONSTMFSLPENSWTPKACVAYY-----IEDDEGELISDVLKPIVQ-----LVF 555
Dh 484 QVREPGQDVLVPLSITTFPIFSRVLVAYYVIGASGREVYABDSVWYDVYDQDSCVSGSLV 543
OY 556 KNKIKLYMSKYKAE-----PSEKYSLRISYTOPDSIVGIVAVDKSVNLMMNSNDITMENY 610
Dh 544 KSG-----QSEDRQCPVPGQOMTLKIEGDHGARV-LVAVDKGVFLINKKNKLTOSKI 594
OY 611 VHELELVNTGYVLGMFMNNSFAVPQECGL 638
Dh 595 WDVVEKADIGCTPGSGKDYAVGVPFDAGL 672

```

Search completed: January 15, 2004, 18:15:19  
Job time : 33 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2004, 18:13:05 ; Search time 44 Seconds  
(without alignments)  
6636.344 Million cell updates/sec

Title: US-10-020-095-4

Perfect score: 7348  
Sequence: 1 MGGPPLITRAHLICVCTAL.....HSSVIFFCFKLLYFMEWL 1428

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 762491 seqs, 204481190 residues

Total number of hits satisfying chosen parameters: 762491

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Published Applications\_AA\*  
2: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
3: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*  
5: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*  
6: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*  
7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*  
9: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*  
17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3360	45.7	665	US-10-108-260A-3396	Sequence 3396, Ap
2	1854.5	25.2	1508	US-10-369-493-5314	Sequence 5314, Ap
3	1850	25.2	1519	US-10-369-493-5313	Sequence 5313, Ap
4	1441	19.6	1474	US-10-292-081A-15	Sequence 15, Appl
5	1441	19.6	1474	US-10-052-817-2	Sequence 2, Appl
6	1440	19.6	1474	US-09-873-403-5	Sequence 5, Appl
7	1440	19.6	1474	US-10-292-081A-10	Sequence 10, Appl
8	1440	19.6	1474	US-10-292-081A-12	Sequence 12, Appl
9	1440	19.6	1474	US-10-292-081A-13	Sequence 13, Appl
10	1440	19.6	1474	US-10-331-496A-38	Sequence 38, Appl
11	1437.5	19.6	1508	US-09-756-247-4	Sequence 4, Appl
12	1433	19.5	1450	US-09-756-247-23	Sequence 23, Appl
13	1433	19.5	1451	US-09-756-247-25	Sequence 25, Appl
14	1428	19.4	1500	US-10-292-081A-9	Sequence 9, Appl
15	1421.5	19.3	1492	US-09-981-151A-10	Sequence 10, Appl

16	1419	19.3	1451	12	US-09-756-247-24	Sequence 24, Appl
17	1384	18.8	1500	12	US-10-316-253-267	Sequence 267, App
18	1384	18.8	1500	12	US-10-316-253-269	Sequence 269, App
19	1255	17.1	1285	9	US-09-925-301-1394	Sequence 1394, Ap
20	1157	15.7	1411	12	US-10-094-886-124	Sequence 124, App
21	1113.5	15.2	941	12	US-09-981-151A-39	Sequence 39, Appl
22	1109.5	15.1	1436	12	US-10-094-886-126	Sequence 126, App
23	1086	14.8	936	12	US-09-981-151A-35	Sequence 35, Appl
24	1080	14.7	898	12	US-09-981-151A-36	Sequence 36, Appl
25	1075	14.6	936	12	US-09-981-151A-37	Sequence 37, Appl
26	1059.5	14.4	931	12	US-09-981-151A-38	Sequence 38, Appl
27	960	13.1	1661	11	US-09-842-758-42	Sequence 42, Appl
28	956	13.0	1663	11	US-09-842-758-43	Sequence 43, Appl
29	934	12.7	1663	11	US-09-842-758-41	Sequence 41, Appl
30	931	12.7	1663	9	US-09-875-519A-22	Sequence 22, Appl
31	921	12.5	1602	9	US-09-778-927A-59	Sequence 59, Appl
32	894	12.2	1642	10	US-09-925-442-2	Sequence 2, Appl
33	880	12.0	1648	10	US-09-925-442-35	Sequence 35, Appl
34	804.5	10.9	751	12	US-09-981-151A-80	Sequence 80, Appl
35	796	10.8	1612	11	US-09-842-758-6	Sequence 6, Appl
36	758.5	10.3	1333	10	US-09-925-442-34	Sequence 34, Appl
37	748	10.2	1251	9	US-09-778-927A-58	Sequence 58, Appl
38	747	10.2	643	12	US-10-292-081A-11	Sequence 11, Appl
39	747	10.2	643	12	US-10-087-188-2	Sequence 2, Appl
40	738	10.0	912	12	US-09-756-247-29	Sequence 29, Appl
41	733.5	10.0	882	12	US-09-756-247-39	Sequence 39, Appl
42	719.5	9.8	528	11	US-09-764-876-20	Sequence 20, Appl
43	718.5	9.8	531	9	US-09-764-853-583	Sequence 583, App
44	689	9.4	134	12	US-10-160-162-139	Sequence 139, App
45	689	9.4	134	12	US-09-820-649-139	Sequence 139, App

## ALIGNMENTS

RESULT 1  
US-10-108-260A-3396  
; Sequence 3396, Application US/10108260A  
; Publication No. US20040005560A1  
; GENERAL INFORMATION:  
; APPLICANT: HELIX RESEARCH INSTITUTE  
; TITLE OF INVENTION: NO. US20040005560A1 full length cDNA  
; FILE REFERENCE: H1-A0106  
; CURRENT APPLICATION NUMBER: US/10/108,260A  
; CURRENT FILING DATE: 2002-03-27  
; NUMBER OF SEQ ID NOS: 548  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3396  
; LENGTH: 665  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-108-260A-3396

Query Match 45.7%; Score 3360; DB 12; Length 665;  
Best Local Similarity 100.0%; Pred. No. 4.9e-268;

Matches 657; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MGGPPLITRAHLICVCTALAVAPGPRILVTPAGIIRPGANTTICVELLEHCPSQVTVYA	60
DB	1	MGGPPLITRAHLICVCTALAVAPGPRILVTPAGIIRPGANTTICVELLEHCPSQVTVYA	60
QY	61	ELTKASNTVTSVLAEGVFEKSKFTLTPLNSADEIYELAVTGTQDEILFNSNT	120
DB	61	ELTKASNTVTSVLAEGVFEKSKFTLTPLNSADEIYELAVTGTQDEILFNSNT	120
QY	121	RLSFTKTSIVFIQTDKALYKPKQKFRIVTLFSDFPYKTSNITLIKDPGNTIQWL	180
DB	121	RLSFTKTSIVFIQTDKALYKPKQKFRIVTLFSDFPYKTSNITLIKDPGNTIQWL	180
QY	181	SOOSDLGVISKTFOULSHPIIGDWSIQVAVDQTYGQFVSEYVLPKEVTLQTPLYCS	240
DB	181	SOOSDLGVISKTFOULSHPIIGDWSIQVAVDQTYGQFVSEYVLPKEVTLQTPLYCS	240

```

Qy 241 MNSHNLNGTITAKTYGKPVKGDVTLTFPLPSFWGKKKNTTKTKINGSANFSNDEEMK 300
Db 241 MNSHNLNGTITAKTYGKPVKGDVTLTFPLPSFWGKKKNTTKTKINGSANFSNDEEMK 300
Qy 301 NMVSSNGLSEYLDLSSPGVEILITTTVESVTGTSRANSTNVFKQHDYIIIEFDYTVL 360
Db 301 NMVSSNGLSEYLDLSSPGVEILITTTVESVTGTSRANSTNVFKQHDYIIIEFDYTVL 360
Qy 361 KPSINFTATVATRADGNQLTLBERNNVVTITQRYNTEYWSGNSNGNQMEAVOKINY 420
Db 361 KPSINFTATVATRADGNQLTLBERNNVVTITQRYNTEYWSGNSNGNQMEAVOKINY 420
Qy 421 TVPOSGTFKIEFPILIEDSSELQKAYFLGSKSSMAVHSLFKSPKTYIQLKTRDENIKVG 480
Db 421 TVPOSGTFKIEFPILIEDSSELQKAYFLGSKSSMAVHSLFKSPKTYIQLKTRDENIKVG 480
Qy 481 SPFELVGNKRLKELSTMVVSRGQVAVGKNGSTMFLTPENSWTPKACVIYVYIEDBG 540
Db 481 SPFELVGNKRLKELSTMVVSRGQVAVGKNGSTMFLTPENSWTPKACVIYVYIEDBG 540
Qy 541 EISDVLKIPVOLVFNKIKLYMSKVKAEPSEKYSLRISTOPDSIVGIVAVDKSVNLMN 600
Db 541 EISDVLKIPVOLVFNKIKLYMSKVKAEPSEKYSLRISTOPDSIVGIVAVDKSVNLMN 600
Qy 601 ASNDITMENVVEHELELYNTGYLLGMFMNSFAVFOEGIMVLTJDTANLTJDKYIDGYDN 657
Db 601 ASNDITMENVVEHELELYNTGYLLGMFMNSFAVFOEGIMVLTJDTANLTJDKYIDGYDN 657

```

# RESULT 2

```

US-10-369-493-5314
; Sequence 5314, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 5314
; LENGTH: 1508
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-369-493-5314

```

Query Match 25.2%; Score 1854.5; DB 12; Length 1508;  
 Best Local Similarity 31.8%; Pred. No. 4.2e-143;  
 Matches 476; Conservative 292; Mismatches 562; Indels 169; Gaps 44;

```

Qy 15 VCTAALAAAPGRPLVTAAGIIRPGANTTIGVELLEHSPQVYKAEILKTAASLTVSVL 74
Db 28 VSTTAAPKPA-TYMLVAPAVRPPQPFVCMNLKQATDEDMVRIEVRTERETJAAR 86
Qy 75 EABGVFEKSGFKTLTLPPLNSAD--EYELRVGTQD--EILFNSSTRLSFETKRLSV 131
Db 87 VISNL-KGIAQYTSLSMPAQSILTPROSYKLYRGETLNAELLFENENELKYQKALSV 145
Qy 132 FIQTDKALYKPKQEVKFRIVTLFSDPKPKYKTSNLILKDKPSNLIQOMLSQOQSLGVISK 191
Db 146 FIQDRAIRPASYIVRYRAIVKSDLKPKYGNATIKIPDSRNLSQITIGVTLDRGVYSG 205
Qy 192 TFOGSSHIIGDMSIQVQ---VNDQTYOSFOVSEYVLPKFEVTLQRPPLXCSMNSKLN 247
Db 206 ELQLAETLLGDWFLVEVTSNGVQDKS---SFTVDYVLPKFEVNIKTSFTITND-DLS 261

```

```

Qy 248 GTTAKTYGKPVKGDVTLTF-LPLSFW-----GKKKN---ITKTKIN--GS 289
Db 262 VFADAKTYGKGVAGKAKYSELPLPMHMAVPTIIDENGVKKEEELMVERTVKLNQGE 321
Qy 290 ANSFNDEEMKNNMDSNGISEYLDLSSPGVEILITTTVESVTGTSRANSTNVFKQHDY 349
Db 322 AAAVFSDNDELK-----HKLHEWGGSIIIVASVEDITEIENNAHQISTFEEV 373
Qy 350 IIEFDYTVTLKPSLNFATVATRADGN--QLTLBERNNVVTITQRYNTEYWSGNS 407
Db 374 KLDVENKGFDFKGLITNVVVALKQMDDTFVKATLPR-----VQSTFFNYVP--YHNDS 427
Qy 408 GNQMEAVQKINTVPOSGTFKIEFPILIEDSSELQKAYFLGSKSSMAVHSLFK----- 461
Db 428 SLQSEKTKIIVEADHGTSVLTQPPINCTSARIEAH-YDIGKDNFTATPISLSLYEA 486
Qy 462 --SPSKTYIQLKTRDEN-IVGSPFELVGNKRLKELSTMVVSRGQV-----AVGKN 513
Db 487 AVSPTKSEIQLADNEADVGVKSLSPSLKATQPLSTITYQVMSRNIIVSSQOMTVNSEH 546
Qy 514 STPESLTPENSWTPKACVIYV-YIEDGELIISDVLKIPVOLVFNKIKLYMSKVKAEPSE 572
Db 547 ATI-SPATANMAPKSLIYALIESQEVVVALDPKVGITQNOVALSIDQAVEPQ 605
Qy 573 KSLRLSVTPQDSIVGIVAVDKSVNLMNASNDITMENVVEHELELYNTGYLLGMF----- 626
Db 606 NVAFKVT-SPKNSFVGLVVDQSVLLKTGNDITREKVEQDLBNYDSNNVGGFGGPRPW 664
Qy 627 -----MNSFAVFOEGIMVLTJDTANLTJDKYIDGYDN-----YIDGYDNA 658
Db 665 EALDRKRSIMRPWWGIGSSDAOSIFSNAGLVLTALLVREFQREFMSVMMWDGAPGA 724
Qy 659 EYAEFEENEENEGHIVDIDHPSLGSS---PHVRKHPETWTIMLDTNNMGYRIYQFEVTPD 715
Db 725 EAA--FAAPMG-----GSSPEPTVRKFPHTWTMSLN-STQSEVMEIEAPD 771
Qy 716 SITSWATGVVISDGLGLTTPVELQAFOPFPIPLNDPSYVIRGEFALEITINYLK 775
Db 772 TITSWASTAINEENGELGVAPTTSKLRVPRPFIQLNLPYAVRGEKFPALLVLFVYME 831
Qy 776 DAIEVKIIT--EKSQKEDILMTSET---NATHQOTLLVPSDGN---VLEPIRPTH 827
Db 832 KEQDVTYLLKXDSGDILKKGDTVVRDEVOQNVRIYVAGGSGSKAVPEPIYESSI 891
Qy 828 GEIPITVTLASPTASDAVNTOMILVKAEGLEKSSOSILDLTNRLOSTLTKTSFSPFPN 887
Db 892 GEIPVHSAISAGGDAVENMLKVDQGYVDNIPVIDLNNSSDPS-KNELIWPND 950
Qy 888 TVTGSERVQITAGDVLGPSINGLASLIRMPYCGCGEONMIFAPNIYILDYLRKKQLTD 947
Db 951 VVDGSGKARLDVIGDMWGPFVLANNAHKLQWPGYCGEONMLNVLNIVKYLRATNRNES 1010
Qy 948 NIKKALSPKQGYQRELYQREDGFSAGNVDPSGTYLSAFVLRCPLEADPYDIDQ 1007
Db 1011 QLETKALKFTEQIQRELYTKRADNSFAGDSKAGSTWLTAFVARSFHAHQYAFVDP 1070
Qy 1008 NVLHRTYTLKGHOKSGEFGWDPGRVYHSELQSG-NKSPVTLTAYVTSLLGKRYQPN 1066
Db 1071 NVISRAVAFINSQOMESGAPFAERGEVHHKDMQGAODGVALLAFVILISL-----ENGM 1125
Qy 1067 DVQESIHFLSEFSRGSIDNYTLALITVALISVGSPPAKKALNMLTWR--AEQEGNQFW 1124
Db 1126 ENKGAVTYLEKHLDEVSGNATMAVVAVALQKKSQAGAFENLKHKGLVKESGDVKA 1185
Qy 1125 VSSE--SKLSDS---WQPSGLDIEVAVALSHFLQFQISSEGIPIYRWLSRQNSLIGF 1178
Db 1186 SAQKVKVEKLESRAYWFOAPVDIETTSYAVLSYLQMNQTSSELSIIRMLVVSQRNBLG 1245
Qy 1179 ASTQDTVALKALSEFAALNTERTNIOQTV--GSPSPS-----PLAVVQPM----- 1224
Db 1246 TSTQDTVALQALSSIAAVYTSKHTSQVITLNGKHTSDINIRKNAVLSQSLSLND 1305
Qy 1225 AVNISANGFGFALCOLNVVYVNAAGSSRRRRRSIQNEAFDLDAVAKENKODINHVLNV 1284

```



Db 1306 AVSINLNGGVFAQLSYST-----YRSLNDAPFCSCQETKEIRAG-NRLQDL 1355  
QY 1285 CTSFSGPGRSGMALMEVNLISGFMVPESEALISLSETVK-----KVEDHGKLNLYLDSVNE 1339  
Db 1356 CCNTRTFGSGNMALAEIDALSGRFADEQVHTLTSIEDQRYEMEDCDTRKMYENPLGG 1415  
QY 1340 TQFCVNIIPAVRNEFKVSNTODASVSIYDYEPERRQAVRSYNSVEKLSCDLCSDVQGRP 1398  
Db 1416 RPYCSTLSYSDVTYQVADQKRPANFRLVDYDPEEQKMTYAAKQTRSLQKCG--EDCWP 1472  
RESULT 3  
US-10-369-493-5313  
Sequence 5313, Application US/10369493  
Publication No. US20030233675A1  
GENERAL INFORMATION:  
APPLICANT: Cao, Yongwei  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
APPLICANT: Goldman, Barry S.  
APPLICANT: Chen, Xianfeng  
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
FILE REFERENCE: 38-10(52052)B  
CURRENT APPLICATION NUMBER: US/10/369,493  
PRIOR FILING DATE: 2003-02-28  
PRIOR APPLICATION NUMBER: US 60/360,039  
NUMBER OF SEQ ID NOS: 47374  
SEQ ID NO 5313  
LENGTH: 1519  
TYPE: PR1  
ORGANISM: Caenorhabditis elegans  
US-10-369-493-5313  
Query Match 25.2%; Score 1850; DB 12; Length 1519;  
Best Local Similarity 31.6%; Pred. No. 9.9e-143;  
Matches 477; Conservative 290; Mismatches 563; Indels 180; Gaps 44;  
QY 15 VCTAALAVAPRPFVLTAGIIRPGNAVITIGVLEHCHSPQVTVKAKLLKTNLSNLTVSVL 74  
Db 28 VSTTAPAVRPA--TYMVAVAVRPPDPFSCVNMLLQATDEDMIVRIEVRTEBNETIAR 86  
QY 75 EAGVGEKSGFKTLT.PSLPLNSAD--EIVELRVNRTQD--EILFENSTRLSFETRIASV 131  
Db 87 VLSNL-KPGIAQVLSLSEMPAOSLTPROSTKLYIRGETNAELIFENNELKTDQKALSV 145  
QY 132 FIOTDRAKYKQKQEVFRITVTLFSDFPKPYKTSINILIKDKPSNLIQOMLSQOSDLGVISK 191  
Db 146 FIOTDRAIYRPAVLAVRAIVKSDLKPYGNATIKIPDSRNLISQITIGVTLDRGVSG 205  
QY 192 TPOLSHPIIDGMSIVQ---VNDQTYOSFOVSEYVLPRKEVTLQTPLYCSMNSKLN 247  
Db 206 EIQLAETTLIGDFIEVETSNGVQDS---SFTVDYVLPKFEVINIKTSSFTIIND-DLS 261  
QY 248 GTTLATYTGKPKPGVTLTF--LPLSPW-----GKKKN-----ITKTPKLN--GS 289  
Db 262 VFIDAKTYTGKGVAGAKVSLPLMRHMAVPTIIDENGVKKEELMVERTVKLRQGE 321  
QY 290 ANFSFNDSEMKVNDWSNGLSEYLDLSSPGVAILTTVESYTGISRNVTNVFFKQHDY 349  
Db 322 AAVVFENDELKR-----HKLLHEMGGSIRIVASVTEDEITERNNTHQISTREEV 373  
QY 350 IIEFPYTVLKPSLNPLFATVKTADGN--QUTLEERNNVITVTOGNYTEYSGNS 407  
Db 374 KLDVERKQGTFFKGLTYNNVVALKQMDPVPVKATLTKR---YQVSTFNYR--YNHDS 427  
QY 408 GNGKMEAVOKINTVQSGTFKIEFPILBDSSELQKAFELSGSKSMAVHSJFK----- 461  
Db 428 SLQEEKETKIVEVDAGTSLVTLQPIINCTSARIEAH-YDIGGKDFKFTPTPYSILVEA 486  
QY 462 --SPSKTYIQLKTRDEN-IKVGSPPELIVVSGNRLKELSMVNVSRQLV-----AVGKON 513

Db 487 AVSPKTSFLQLADNEGAVDVGKSLSPSLKATQPLSTITTYQVMSRNVIVSQOMTVNSEH 546  
QY 514 STMFLTPENGSMTPACIYV-YTEDGEIISDVLIKIVQVLPKKIKLYMSKYAESE 572  
Db 547 KTI-SFPAITAMPKSRILITVAILSSQEVLDALDFVEBGFQVQVALSIDKQAVEBQ 605  
QY 573 KVSLSISVTPDSIYGVAVDVSVNLMMNASNDITMENVHELELYNTGYLGMP----- 626  
Db 606 NKPKFVY--SDKNSFVGLLVVQSVLLTKTGNDITREKVEQDLENDSNNVGGFCGPRPW 664  
QY 627 -----MNSFAVQEGGLMVLTD-----NLT 647  
Db 665 EALDKKESIMRPMWIGSGSDAQSIFFSNAGLVLLTDLALYRPPQREFMSERRLTPGLLT 724  
QY 648 KDYIDGVYDNAMEYARFMEENEGHVDLHDSLGS---PHYRKHFPETWIMLDNMGCR 704  
Db 725 VMMMDGACGMABAA--FAAPMG-----GSSPPPTVRKFFPHTWIMSDLN-STS 771  
QY 705 IYQEFVTPDPSITSMVATGFVISEDLAGLTTTPVELQAFQPFIFPLNLYSVYRGEEF 764  
Db 772 GEVMEIEAPDITISVASTPAINEBGLGVAPTTSKLRVRRPFIQNLTYAARRGKF 831  
QY 765 ALKITIFNYLDATVEKVII--EKSDKFDILMTSSEI--NATGHQOTLVPSEDAF-- 817  
Db 832 ALLVAVFVYMEKEQDVTVTLKDYKDSGYDLKKQGTVRDEVGQVNRIVSVAGGTSK 891  
QY 818 -VLEPIRTHLGEIPITVYALSPTASDAVYQOMILYKAGIEKYSQSITLLDITNRLQST 876  
Db 892 AVEFPVIVSSIGELPVHISALASQGGDAVENKLRDPQGYVDNRIPVITDINNNSDPS 951  
QY 877 LKTLSPSPPTVVTSSERVOITAIQDVLGSPINGLASLIRMPYGGGEONMIFANITYIL 936  
Db 952 -KNLELIPNDVVDGSCARLDVIGDMGPVLNNAHKLVOHPYGGGEONMLNVLNITLV 1010  
QY 937 DYLTKKQDITNLKELKALSPKQGYORELLYQREDSFSAFGAYDPSGTYLSAFVIRCF 996  
Db 1011 KILRATNRNESQLETKAIKFIEQIGRELTYKRADNSFSARGSDSKASTWLTAVVASF 1070  
QY 997 LEADYIIDIQNVLTARITVTLKGHQKSGEPMWDERVTVHSELQCG--KNSPTLTAYITS 1055  
Db 1071 HHAKQYAVDPVPSVSRVAFVNSQOMESGAFAEREVHHKMDQGGADGVALTAFVLIS 1130  
QY 1056 LLGYRKYQPDIDVQSHIFLSEPSRGISDNVYTLALITYALSVSQSPKAKELMVLTR- 1114  
Db 1131 IL-----ENGEMNGKAVITYLEKHLDVSGNAYTMVAVALQAKSKQAGAFENLKXHK 1185  
QY 1115 -AEQEGNQFVWVSE--SKLSDS---WQPSRLIEVAVALLSHFIQFQTSBGPIMRW 1167  
Db 1186 IVEKSGDVKFAQAQKVEKLKESRAYMFOARPVIDETTSYAVLSLAQNGTSELSITRW 1245  
QY 1168 LSRQNSLGGFASITDITVALKALSEFALNTERNTIQTV--TGPSSPS-----PLAV 1220  
Db 1246 LVSCNNEIGEGFSTODTVMALQALSSVAAVYVSDKHSQVITLNGKHTHSIDINRNAIV 1305  
QY 1221 VQPM-----AVNISANGFGALICQALNVVNVVAVASGSSRRRBSIQONEAFDLDAVVEN 1273  
Db 1306 LQSYLSSLNDVAVSINANGTVVFAQLSYST-----YRSLNDAPFCSCQETKEI 1356  
QY 1274 KDLNHLVDLNVCTSPSGRSGMALMEVNLISGFMVPESEALISLSETVK-----KVEYDHG 1328  
Db 1357 RAG-NRLQDLCCNTRFGSKNMALAEIDALSGRFADEQVHTLTSIEDQRYEMEKD 1415  
QY 1329 KLNLYLDSVNETQFCVNIIPAVRNEFKVSNTODASVSIYDYEPERRQAVRSYNSVEKLS 1388  
Db 1416 KKNVVFENPLGGRPVCLSYSDVTYQVADQKRPANFRLVDYDPEEQKMTYAAKQTRSLQ 1475  
QY 1389 LQSDVQGRP 1398  
Db 1476 KCG--EDCWP 1483  
RESULT 4

US-10-292-081A-15  
 ; Sequence 15, Application US/10292081A  
 ; Publication No. US20030162202A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kenneth David Becker  
 ; APPLICANT: Gonul Velicelebi  
 ; APPLICANT: Xin Wang  
 ; APPLICANT: Randolph E. Tanzi  
 ; APPLICANT: Lars Bertram  
 ; APPLICANT: Aleister J. Saunders  
 ; TITLE OF INVENTION: SINGLE NUCLEOTIDE POLYMORPHISMS AND MUTATIONS ON ALPHA-2-MACROGLOB  
 ; FILE REFERENCE: 37481-3323  
 ; CURRENT APPLICATION NUMBER: US/10/292, 081A  
 ; CURRENT FILING DATE: 2002-11-08  
 ; PRIOR APPLICATION NUMBER: 60/337434  
 ; PRIOR FILING DATE: 2001-11-09  
 ; NUMBER OF SEQ ID NOS: 15  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 15  
 ; LENGTH: 1474  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-10-292-081A-15

Query Match 19.6%; Score 1441; DB 12; Length 1474;  
 Best Local Similarity 29.2%; Pred. No. 5.5e-109;  
 Matches 453; Conservative 264; Mismatches 551; Indels 284; Gaps 60;

QY 12 LCCCTALANAPGRFVLTAPGIIIRPGNAVITGVLELHPCSOVYVYAKELKTAASLTV 71  
 DB 15 LVLVLPDASVSGKQYVAVPSLHT-ETTEKGVLLSYLNETVYVYASLESVGRKSL 73  
 QY 72 -SVLAEG-VFEKSGFKTLTLPSPINSGADE---IYELRVTRQDELTFNSNTRLSPET 126  
 DB 74 FTDLAENDVLHCAAF-----AVPKSSNEBWFVTVQVGPQOE---FKKRTTWAKN 124  
 QY 127 KRISVFIQDRAKYKQEVKFRIVTLFSDFKPKYKTSNLIL-IKDPKSNLIQWLSQSD 185  
 DB 125 EDSLVFVQDTSIKYKPGQTVKFRVYVMDENFHPNLNELPLVYIDPKRNIAQMSFQLE 184  
 QY 186 LGVISTKQLSHPLPGLDMSIQVQVND--QTYVQSFQVSEVVLKPEFVTLQTPYCSNNS 243  
 DB 185 GGLKQFSPPLSEPPQSGYKVVQKSGGRTGH-PFYVEFVLPRFEVQVTPPKITITILE 243  
 QY 244 KHLNGTIFAKTYGKPVGDTLPLPL-----SFMGK 276  
 DB 244 EEMAVSVGGLTYTGKPVGHVTVSICRKYSDASDCHGDSQAFCEKBSGQLNSHCCTYQ 303  
 QY 277 KKNITKTFKINGSANFSFNDEMNKVMDSNGLSEYLDLSSPGVEIIT-----TVTESV 331  
 DB 304 VK--TKVQQLR-----KEYEMK-----LHTEAQIQEGTVVELTGRQSEIRRTI 347  
 QY 332 TGISRNVSTNVFPGKHDIIEFPDYTVLKPSLNFATATVKTTRADGNOLTEERRNNVYI 391  
 DB 348 TKLS-FVVDVSHFRG--IPEFG-----QVRLVDGKGVPIP--NKVI 384  
 QY 392 TVTQRNTEYVSGNSGNGQKMAVO-KINYT--VPOSGTFKIEPP-----LLEDS 438  
 DB 385 FI-RGNENANY--SNATTDHGLVQFSINTNNMGSTLTVRVNKKDSSPCGYMVSEBH 441  
 QY 439 SELQKAFYFLGSKSMVAHSLFKSPSKTYIQLKTRDENIKVGS---PFLVVSNGK--R 492  
 DB 442 EEAHTTA-----LVFSPSKSFVHLERPSHSLPCGHQTVQAHITLNGTILG 489  
 QY 493 LKELSL--YMAVSRQGLVAVG-----KQNST---MPSLTPENSWTPKACVITYIEDDG 540  
 DB 490 LKKSIFYLLIMAKGIVATGTGILLVQEDMKGHFSISIPKSDIAPARLLIYAVLPTG 549  
 QY 541 EIIIDVLKIPVQVLFKKNIKLYMSKVAPESEKSLRISTVQPSIYIGIVAVDSVNLML- 599  
 DB 550 DVIDSDAAYDVENCILANKVDLSFSPSOSLPASHHLRLRYTA--POSVCLARVVDSDVLLMK 608  
 QY 600 ----NNS-----NDITMENVVHLELYNTGY--YLGMFMNS-----FA 631

DB 609 PDALSSASVYNNLPEKDLTGFPGLNDQDNEDCINRHNHYINGITYTTPVSTNEKDMYS 668  
 QY 632 VFOECGLMVLTDANLTQDYIDGYDNAE-----YAEFMENECHIIDHDFSL 680  
 DB 669 FLEDMGLKAFNSKIRKPKKCPQLQOYEMHGPGLKGVFYESVWGRHARLVHVE-- 725  
 QY 681 GSSPH--YKHPPEMTWIDTMGYRIYOEFEVTVDSITTSVWATGAVLSEDLGLT 737  
 DB 726 ---PHTEVAKYFPEMTWIDLVVNSAGVAVCVTPDITTEKKAFCLESDAGCIS 782  
 QY 738 TPVELAQFPFPLNLPYSVINGEPALEITFNILKATEVKVILEKDKDILMTSS 797  
 DB 783 T-ASLRAFQFPFELTMPYSVIRGEAFTLKATYANLTPKICIRSVOLDSAPLAPEVK 841  
 QY 798 E-----INATGHQOTLLVPESDAGVLEFPIRPHLEGIPTTVYALS-----PTA 841  
 DB 842 EQAPHICANGRO-----TVSMVTPSLSGVNTVVAEALESOLCTGTEPSV 890  
 QY 842 ----SDAVTQMLVVAEGIEKSYOSILLDLTDNRLOSTLTKTSPSPNTVTSERVQ 896  
 DB 891 PEHGRKOTVAKPLVPEEGLEKETTSNLCPGSGEVS---ELSLKLPNVVVEESARAS 947  
 QY 897 ITAIGDVLGPSINGLASLIMPYGCEBQNMNINAPNIYIIDYITKKKOLTDNLKERALS 956  
 DB 948 VSVLGDILGSAMONTQNLQMPYGCGBQNVLPAPINIVYVDINETQOLTPREIKSKAIGY 1007  
 QY 957 MRQYORELLYOREDSFSAFG--NYDPS--GSTWLSAFVLRCELEADPYIDQNVLHRTY 1014  
 DB 1008 LMTGQKQLNYKHYDGSYTFGERYGRNQNTVLTAVLTFEQALAPYFIDAHITQAL 1067  
 QY 1015 TWLKGHQSNGEFPWDPGRVHSELQGNKSPVLTAVTVTSLSGYRKYQPNIDVQESHP 1074  
 DB 1068 IWSLQKQKDCFRSSGSLNNNAIKGVEDEVLSAVITLALLEIPPLTVHPVYRNALFC 1127  
 QY 1075 LES-----EFSRGISDNTYLLITALSVGS--PRAKEALNLTWRBEOGGMQV--- 1124  
 DB 1128 LESAMKTAQEGDHG--SHVYTKALLAFAFALAGQODRKEVLSLNEBAVKKDSSVHERP 1186  
 QY 1125 VSESKLSDSMOPR--SLDIEVAAYLLSHFLQFO--TSEGI-P---IKRWLSRQNSL 1175  
 DB 1187 OKRPAPGHFPEQDAPBAEEMTSYVLLA-YLQAQPAPISEDLTSATNIVKMITKQOQAQ 1245  
 QY 1176 GFAPSTQDTTVALAKALSEFALANTRNTNIQVTVTGSPSPPLAVVO----- 1222  
 DB 1246 GFSSSTQDTTVALHALSKYAATFT--RTGKAQVTTIGSSGTFSSKKQVDDNNRLLIQVS 1304  
 QY 1223 ----PMAVNISANGFPAICQLNVTVNVKASGSSRRRSRTIQOEAFLDVAVK--ENKD 1275  
 DB 1305 LPELPGESYMKVYEGCVYLTSLAKYNI-----LPEKEEPFALGVQTLPTQCD 1353  
 QY 1276 DLN-HYDLANCTSFSGG---RSGMALMEVNLISGFVPEBALST---SFTVKVYEDHG 1328  
 DB 1354 EPRKHTSFQISLSVSTGSSASNMALVDKVMVSGTPLKPYTKMLERSHVSRTESVN 1413  
 QY 1329 KNLVYDSVNETQFCVNIIPAVRNFKVSNQDASVSIYDYVEPRQAVRSYS 1380  
 DB 1414 HVLITDKVSNQTLSLFFTVLQDVPVRDLKPAIVKYDYETDEFALAEYNA 1465

RESULT 5  
 US-10-052-817-2  
 ; Sequence 2, Application US/10052817  
 ; Publication No. US20020114792A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Tanzi, Rudolph E.  
 ; APPLICANT: Kovacs, Dora  
 ; APPLICANT: Saunders, Aleister J.  
 ; TITLE OF INVENTION: Alpha-2-Macroglobulin Therapies and Drug Screening Methods for  
 ; FILE REFERENCE: 0609, 4460005  
 ; CURRENT APPLICATION NUMBER: US/10/052, 817  
 ; CURRENT FILING DATE: 2002-01-23

; PRIOR APPLICATION NUMBER: 09/241,606  
 ; PRIOR FILING DATE: 1999-02-02  
 ; PRIOR APPLICATION NUMBER: 09/148,503  
 ; PRIOR FILING DATE: 1998-09-04  
 ; PRIOR APPLICATION NUMBER: 60/093,297  
 ; PRIOR FILING DATE: 1998-07-17  
 ; NUMBER OF SEQ ID NOS: 27  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 2  
 ; LENGTH: 1474  
 ; TYPE: prt  
 ; ORGANISM: Homo sapiens  
 US-10-052-817-2

Query Match 19.6%; Score 1441; DB 14; Length 1474;  
 Best Local Similarity 29.1%; Pred. No. 5.5e-109;  
 Matches 451; Conservative 266; Mismatches 551; Indels 284; Gaps 59;

QY 12 LILCVTAALAVAGPRELVTAPGIRPGNVTIGVLEHSCPSQVTVKALKLTASNLTV 71  
 DB 15 LILVLPPTDASVSGKQYQVWLVPVSLHT-ETTEKGCVLISLNEVTVASLESVRGNRL 73  
 QY 72 -SVLAEAG-VPEKSGKTLTLPVSLNSADE--IYELAVTGRTODEILFNSSTRLSFT 126  
 DB 74 FTDLKENDVILHCVAF-----AVPKSSNEEVMFLVQVKGPTQE--FKKRTVVMVKN 124  
 QY 127 KRIASVFIOTDKALYKPKOEKFRIVTLFSDPKYKTSNLIL-IKPKSNLIOQMLSQOSD 185  
 DB 125 EBSLVFVQIDKSIYKFGQTVKRVVSMDEKHFHLELVLVYIOPKRNLIQWOSFOLE 184  
 QY 186 LCVISKTFQLSHPILIGDMSIOVQVND--QTYVOSFOVSEYVLPKFEVTLQTPLYCSMNS 243  
 DB 185 GGLKQFEPFLSEPFQSGYKVVQKSGGRTEH-PTVESEVLPKFEVTVTPKTIITILE 243  
 QY 244 KHLNGITTKYTYGKPVKQDVTLLPL-----SFMCK 276  
 DB 244 EEMNVSVCGLYTYGKVPQHVTVSICRKYSDASDCHGEDSOAFCEKFGQLSHGCFYQ 303  
 QY 277 KKNITFTKINGSANFSPNDEEMKVNWSSNGISEYLDLSPGVEILT-----TVTSBV 331  
 DB 304 VK--TKVFLK-----KEYEMK-----LHTEAQIOEGTVVELGROSSSEITRTI 347  
 QY 332 TGISRNVSTNVFEKQHDYIIEFPDYTYTLKPSLNFATVAVKTRADGNQTLLEERNVYI 391  
 DB 348 TKLS-FVKVDSHPRQ--LPPFG-----QVRLVQKGPPI-----NKYI 384  
 QY 392 TYTORNYTEYSGNSGNQKMAVQ-KINYT--VPOSGTFKIEP-----ILEDS 438  
 DB 385 FI-RGNEANY--SNATTDHGLVQPSINTVWVGSLTVRVVYKDRSPCYGVQWVSEH 441  
 QY 439 SELQKAYPLGSKSMAVHSLFKSPSKTYIOLKTNDENIKVS-----PELVVSGNK--R 492  
 DB 442 BEAHHTAY-----LVPSBSKSVHLEPMSHELPCGHQTQVQAHYILNGTLLG 489  
 QY 493 LKELS--YVWVRGQLVAVG-----KONST-----MFSLTPENSMTPKACVIVYIEDDG 540  
 DB 490 LKKLSFYVILMAKGGVTRGTHGLLVKQDMKGPISISIPVSDIAPARALLIYAVLPFG 549  
 QY 541 EIIISVLTLPVQLVFNKIKILYMSKYKAPSSKVSILISVTPDSDIVGIVAVDKSVNLN 600  
 DB 550 DVIQSAKDVENCILANKVDLSFSPQSLPASHAHLRYAA-PQSCALRAVDQSLMLK 608  
 QY 601 ASNDITMEVNHLEL-ELVNTGY-----YLGMFNS-----FA 631  
 DB 609 PAELISSASSVYVNLPEKDLTGFPGLNDODECINRHHVYINGITTYTPVSSSTNEKMS 668  
 QY 632 VFOECGLAWLTDLANTLKDYIDGYDNAE-----YARFMEENGHIVDIHDFSL 680  
 DB 669 FLEDMGLKAFNISKIKPKMCPQLOQYEMHGPBGLRVGVEEDVMGRGAKLVHVE--- 725  
 QY 681 GSSPH--VRKHPEETWLDITNMGRYIOEPEVTVPDSITSWVATGVVISDLGLTT 737  
 DB 726 ---PHTETVRKFFPETWLDLVVNSAGVAVGVTVPTDITTEWKAQAFCLSEDAIGISS 782

QY 738 TPVELQAFQPEFPIPLNTPSVYIRGSEFALETITRNLYKDATEVYVITIEKSDKPDILMTSS 797  
 DB 763 T-ASLARQPFVELTWPYVSIVIRGEAFILKATVNLVLPKIRVSGVQLEASAPFLAVPEK 841  
 QY 798 E-----INATHQOCTLVPSDEGATVLEPIPTHIGELPIVTYALS-----PTA 841  
 DB 842 EAPHCICANGRQ-----TVSWAVPKSIGVNFVSAEALESOELCGTEVPBV 890  
 QY 842 -----SDAVQMLIVKAGIEKYSQSLILDLNLRLOSTLTKTISFSPRPTVYSGEVQ 896  
 DB 891 PEHGRKDTVIRPLVPEBGLKEKTEFNSLLCPSGGEVS--EELSLKLPVNVESABAS 947  
 QY 897 ITAIGDVLGPEISINGLASIRMPYCGEGEQNMINFAPNIYILDYLTREKQTLNLEKALS 956  
 DB 948 VSLVGLDILGAMQNTQNLQMPYCGEGQNMVLPAPNIYILDYLMETQQLTEIKSKALCY 1007  
 QY 957 MRQGOEELLYQREDEGFSFAFG-NYDS-GSTWLSAPVLRCEFLADPYIDIDQVLAHRTY 1014  
 DB 1008 LNTGYQRLNXYKHDXSYSTFGEERYGNQNTWLTAFVLTQFAQARAVIFIDEAHITQAL 1067  
 QY 1015 TWLKGHOKSNEBFPDQGVHISELOGKNSPVTLTAVYVSLGVRKQVPIQVDESIFH 1074  
 DB 1068 TWLQROKQKNCFFSSGSLNNNAIKGVEDEVTLISAYITLALBEIPLTVTHPVVKNALFC 1127  
 QY 1075 LES-----EFSRGISDNYTLALITYALSVSQS-PKAKEALNMTWRAQEGQVFW--- 1124  
 DB 1128 LESAMKTOEEDHG-SHYTTALALAYAPALAGNDKREVYKSLINEAVKQNSVHMERP 1186  
 QY 1125 VSSBSKSDSWQPR-SLIDIEVAAYALLSHFLQO--TSBGP--IMKLSRQRNRL 1175  
 DB 1187 QKRAPVGHFYEPQPSAEVEMTSYVLLA-YLTQAPAPTSBDLISATNIVKMTQOQAQ 1245  
 QY 1176 GGFASDTOTYALKALSFPALNTERNIQVYTYGSSPEPLAVQO----- 1222  
 DB 1246 GGFSSDTOTYVALLHSLKYGATFT-RTGKAQVYTIOSGTFSSPFQVNNRRLLOQVS 1304  
 QY 1223 ---PMAVINISANGGFALCOLNVVNVKASGSSRRRSIONOEAFDLDVAVK--ENKD 1275  
 DB 1305 LPELPEBYSMTYBEGCYLOTLSTKNI-----LBEKEFPALAGVQTLPOICD 1353  
 QY 1276 DLN-HVDLNVCTSPSGP--RSGMALMEVNLGFWPSEBAISL--SETVKKVEYDHG 1328  
 DB 1354 EPKATHSFQISLSTSYTSSRSASNNAIYDVWVGGFILPKETVKKMLERSNHSVREVS 1413  
 QY 1329 KNLVLDVNETQFCVNIIPAVNRKVSNTQDASIVYDYEPKQAVASYS 1380  
 DB 1414 HVLVYLDKVSQTLSPFTVLAQDVVRLKPAIVKYVDYETDEFAIAEVNA 1465

RESULT 6  
 US-09-873-403-5  
 ; Sequence 5, Application US/09873403  
 ; Patent No. US2002028207A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Sistrava, Pramod K  
 ; TITLE OF INVENTION: COMPLEXES OF ALPHA (2) MACROGLOBULIN AND ANTIGENIC  
 ; FILE REFERENCE: 8449-178  
 ; CURRENT APPLICATION NUMBER: US/09/873,403  
 ; CURRENT FILING DATE: 2001-06-04  
 ; PRIOR APPLICATION NUMBER: 09/625,139  
 ; PRIOR FILING DATE: 2000-07-25  
 ; PRIOR APPLICATION NUMBER: 60/209,266  
 ; NUMBER OF SEQ ID NOS: 5  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 5  
 ; LENGTH: 1474  
 ; TYPE: prt  
 ; ORGANISM: Homo sapiens  
 US-09-873-403-5

Query Match 19.6%; Score 1440; DB 9; Length 1474;  
 Best Local Similarity 29.1%; Pred. No. 6.7e-109;  
 Matches 451; Conservative 266; Mismatches 551; Indels 284; Gaps 59;

12 LVCCTAALAVADPFPFLVTAPGIIIRPGGNTTIGVLELHCSPQVYKAEELKTASNLTV 71  
 15 LVLVLPDASVSGKQYVNLVPSLHT-ETTEKGCVLSTYLNETVYASLESVGRNRL 73  
 72 -SVLEABG-VFEKSGFKTLTLPSPLSNLSADE---IYELRYGTQDEILFNSNTRLSFET 126  
 74 FTDLAENDVLHCVAF-----AVPKSSNEEWMELTVQVKGPTQD---FKKRTTVWVKN 124  
 127 KRISVICTDALKYKPKQEVKFRVTLFSDPKPKYKTSNLIT-IPDKSNLIQMLSQOSD 185  
 125 EDSLTFVQTDKSIYKPGQTVKFRVYVMDENFHLNELIPLVYIQPKNRIAQMSFQLE 184  
 186 LGVISTKQSLSHPLTGDMSIQOVND--QTYQSGFOVSEVVLPRFVTLQTPLYCSNNS 243  
 185 GGLQKQSPSLSEPPQSGYKVVQKSGGRTGH-PFTVEFVLPRFEVQTVPKITITILE 243  
 244 KHLNGTIAKTYGKPRKGDVTLTFLPL-----SFMKG 276  
 244 EEMVSVSGCLTYTGKPRVGHGHTVSI CRKYSASDCHGEDSOAFCEKBSGQLNSHGCTYQ 303  
 277 KKNITKTFKINGSANFSFNDKMNWDSNGLSEYLDLSSPGVEYILT---TVTESV 331  
 304 VK--TKVFOLR-----KEYEMK-----LHTEAQIQEGTVVELTGRQSEITRTI 347  
 332 TGISRNSTNVFKQHDITIEFPDVTYKLSLNFATVKTTRADGNOLTEERRNNVY 391  
 348 TKLS-FYVVDHFRQG---IPFFG-----QVRLVDGKGVPIP---MKVI 384  
 392 TVTQRYNTEYWSGNSGQKMEAVQ-KINYT--VPOSGTFKLEPP-----ILEDS 438  
 385 FI-KGNENANY--SNATTDHGLVQFSINTNNWGTSLTVAVNKKDSPCVQWVSEBH 441  
 439 SELQKAYFLGSKSMVAHSLFKSPSKTYIQKTRDENIKVGS---PFEIVSNGK--R 492  
 442 EEAHTTAY-----LVFSPSKSFVHLPEMSHELPCGHGTQVQAHYLLGTTLG 489  
 493 LKEHS--YMWVSROGLVAVG-----KONST---MPSLTGENSKTPKACIVYIYEDG 540  
 490 LKKLSFYLLIMAKGDIVRTGHLGLVQEDMKHFSISIPKSIDAPARLLIYAVLPTG 549  
 541 EISDVLKIPVOLFKNKIKLYMSKVKAEPSEKYSLSISTVOPDSIYVIGVAVDSKVLNM 600  
 550 DVIQDSAKYDVNCLANKVDSLSPSSQSLPASHAHLRTAA-POSCALRAVDSVLMM 608  
 601 ASNDITMENVHEL-ELYNTGY-----YLGFMFMS-----FA 631  
 609 PDAELASASVYNLPEKDLTGFPGLPDLNDQDDECI RNNAVYINGITYPVASSTNEKOMYS 668  
 632 VPOEGGLVLTDAULTKYIDGVIDNAE-----YAEFMEENEGHIVIHPSL 680  
 669 FLEWGLKAFNNSKIRKPKMCPOLQOQYEMHGPBGLRVGFYSDVWGRHARLVHER--- 725  
 681 GSSPH---VRHGFPTWIMDLTNMGYRYOEFEVTVPDPSISWATGVISDGLGLTT 737  
 726 ---PHTETVAKYFPTWIMDLVNVNAGVAVTVPDITTEMKAGAFCSDDGLGIS 782  
 738 TPVELQAFQPFIFLNLVSVIRGEFALETITFNLYLKDATEVYVILIEKSDKFDILMTSS 797  
 783 T-ASLRARQPFVELTWPYSVIRGEAFLLKATVNLVYKPCIRVSGQLEASAPFLAVPREK 841  
 798 E-----INATHGQTLVPSDEGATVLPFIRPMLGELPITVTALS-----PTA 841  
 842 EADPHCICANGRO-----TVSVAVTPKSIKGNVFTVSAABLESQELCTEVPV 890  
 842 ---SDAVTMIIYKAGIEKYSQSILDLTLNRLQSTLKTLSFSPPTVTVGSRVQ 896  
 891 PEHGRKDVLPKLVPEBGLKETTFNSLLCPSSGGEVS--EELSLKIPPVNVEESAPAS 947  
 897 ITAIGDVLGSPINGLASLIRMPYGGGEQNMINFANIIYLDYLTKKQLTDLNLEKALS 956

DB 948 VSVLGDILGSAAMONTONLQMPYGCBOHNVLPAPNIVYLDVLTNETQQLTPEVKSRAIGY 1007  
 QY 957 MRQYORELLYOREDSFSAFG-NYDPS-GSTWLSFVLRCPLEADPYDIDONVHRTY 1014  
 DB 1008 LMTGQKQNLNKKYKDDSYSTFGERYGRNQONTLTFVLKTFQAAAYIFIDEAHITQAL 1067  
 QY 1015 TWLKGHKSNGEFPWDPGRVHSELQGNKSPVTLTAAYVTSLLGYRRKYOPNIDQESHP 1074  
 DB 1068 IWLSPQKQKGRSSGSLNNNAIKGVEDEVLTSAVITALEIPLTVTHPVVRNALFC 1127  
 QY 1075 LES-----EFSKCIDNTLALITYALSVS-PAKELNMLTRABEGEMQW--- 1124  
 DB 1128 LESAMKTAQGDHG-SHVYTKALAAFAFALAGQDRKEVLSKINEAVKDSVHERP 1186  
 QY 1125 VSEESKIDSQMR--SLDIEVAAYLLSHFLQFO---TSEGIP-----IWRMLSRORNSL 1175  
 DB 1187 QKRAKPVGHFEYEQASAEVEMTSYVLLA-YLTAQAPFSEDUTSATNIVKMITKQNAQ 1245  
 QY 1176 GGFASDTDTVALKALSEFALMNTERTNIQVTVTSPSPPLAVVQ----- 1222  
 DB 1246 GGFSSDTDTVALHALSKGAATFT-RTGKAQVTTIQSGTFSSKQVDNNRNLQOVS 1304  
 QY 1223 ---PMAVNISANGFPAITQNLVNVYVYKASGSSRRRSIQNOEAPDLVAVK---ENKD 1275  
 DB 1305 LPELPGESMKVYGEQVYLTQSLKYN-----LPEKEFPFPAIGVQTLPTQCD 1353  
 QY 1276 DLN-HVDLWNTGSPSGP---RSGMALMEVNLISGFVWSEALST---SETVKVEYDHG 1328  
 DB 1354 EPKAHTSFQISLSVSTYSGRSASNMALVDKMGSGFPLKPYKMLERSNHSRTVESSN 1413  
 QY 1329 KLMYLDVSNETQFCVNIIPAVRNFKVSNTOASVSIVDYEPERRQAVRSYNS 1380  
 DB 1414 HVLITLDKVSNGQLSLPFTVLQDVPVRLKPAIVKYDYDETDEFALAEYNA 1465

RESULT 7  
 US-10-292-081A-10  
 ; Sequence 10, Application US/10292081A  
 ; Publication No. US20030162202A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kenneth David Becker  
 ; APPLICANT: Gonul Velliclebi  
 ; APPLICANT: Xin Wang  
 ; APPLICANT: Randolph E. Tanzi  
 ; APPLICANT: Lars Bettram  
 ; APPLICANT: Aleister J. Saunders  
 ; TITLE OF INVENTION: SINGLE NUCLEOTIDE POLYMORPHISMS AND MUTATIONS ON ALPHA-2-MACROGLO  
 ; FILE REFERENCE: 37481-3323  
 ; CURRENT APPLICATION NUMBER: US/10/292, 081A  
 ; CURRENT FILING DATE: 2002-11-08  
 ; PRIOR APPLICATION NUMBER: 60/337434  
 ; PRIOR FILING DATE: 2001-11-09  
 ; NUMBER OF SEQ ID NOS: 15  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 10  
 ; LENGTH: 1474  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-10-292-081A-10

Query Match 19.6%; Score 1440; DB 12; Length 1474;  
 Best Local Similarity 29.1%; Pred. No. 6.7e-109;  
 Matches 451; Conservative 266; Mismatches 551; Indels 284; Gaps 59;

12 LVCCTAALAVADPFPFLVTAPGIIIRPGGNTTIGVLELHCSPQVYKAEELKTASNLTV 71  
 15 LVLVLPDASVSGKQYVNLVPSLHT-ETTEKGCVLSTYLNETVYASLESVGRNRL 73  
 72 -SVLEABG-VFEKSGFKTLTLPSPLSNLSADE---IYELRYGTQDEILFNSNTRLSFET 126  
 74 FTDLAENDVLHCVAF-----AVPKSSNEEWMELTVQVKGPTQD---FKKRTTVWVKN 124

```

QY 127 KRISVFIOTDKALYKPKQEVKFRIVTLFSDFKPYTSLNLI- IKOPKSNLIQOMLSQSD 185
DB 125 EDSLVEFVOTDKSIYIRGQTVKFRVVSMDENFHLNELLPLVYIOPDKNRIAQWOSFOLE 184
QY 186 LGVISTKFTOLSSHPLIDGMSIOVOND--QTYOSFOVSEVYLPRPEVTLQTPLYCSMNS 243
DB 185 GGLKQSPFLSSSEPPQSGYKVVQKSGGRTEH-PTVEEVLPRPEVQVTPKIIITILE 243
QY 244 KHLNGTITAKTYTKPKVKGDTVLTLPL-----SFMGK 276
DB 244 EEMNVSVGLTYTKPKVPGHVTYSICRKYSDASDCHGEDSOAFCEKFSQQLNSHCIFYQ 303
QY 277 KKNITKTKINGSANFSFDEEMKVMSSNGSLSEYLDISPGVEYLTL---TVTSV 331
DB 304 VK--TKVFKLR-----KEYEMK-----LHTEAQIOEGTVLTLGRQSEITRTI 347
QY 332 TGISNVSTNVFKQHDYIIEFDYTVLAKPSLNTATYKTVRADGNQTLLEERNVYI 391
DB 348 TKLS-FVKVDSHFQO---IPFG-----QVRLVQKGVPIP---NKVI 384
QY 392 TVTQNTLEYWGSNSGNOKEAVO-KINYT--VFQSGTKIEFP-----ILED 438
DB 395 FI-RGNEANY--SNATDEHGLVQPSINTNMGTSLTVRNVYKDRSPCYQWVSEH 441
QY 439 SELQKAYFLGSKSSMAVSLFKSPSKTYIOLKTRDENIKVGS---PELVVSGNK--R 492
DB 442 EEAHHTAY-----LVFSPSKSVHLHPMSHELPGHQTQVQAHYIYINGTLLG 489
QY 493 LKELS--YVVSRGOLVAVG-----KQNST---MFSITPNSMTPKACVIVYIETDDG 540
DB 490 LKKSFPYLLIMAKGQIVRGTHGLLVKQEDMKHFSISIPVKSADLAPARLIIYAVLPTG 549
QY 541 EIIISVULKIPLYOVLPKNIKLYWSKVKAPESEKSLRISVOPDSIYGVVANDKSVNLMN 600
DB 550 DVIGSADYDENCANLKVDSFSPQSOLPASHALIKRYA--PQSCALRAVDQSVLMLK 608
QY 601 ASNDITMENVHEL-ELYNTGY-----YLGMFNS-----FA 631
DB 609 PDAELASASVYVNLLEKOLTGFPPLNQDDEDCINRNVYINGITYPVSTNEKDMVS 668
QY 632 VFQSGGLVLDANLTKOYIDGVYNAE-----YAEFNEEGHIVDIHPSL 680
DB 669 FLEDGLKAFNTSKIRKPKMCPQLOQYEMHGPEGLRVGVESDVMGRGARLVHEE--- 725
QY 681 GSSPH---VRKHPEPTWIMLDTNMGYRTYQEFVYVPSIISVAVTGFYISDLGLTT 737
DB 726 ---PTEETVRKTFPEBTWIMDLVVVNSAGVAVGVTPDITTEWKAGACLSDEAGLIS 782
QY 738 TPVELQAFOPPEFPLNLPYSVIRGEFALEITIENYLKDATEVKYIIEKSDFKDIAMTS 797
DB 783 T-ASIRARQPPFVELLMPYSVIRGEAFILKATVANYLPEKICRVSOQLASPAFLAVPEK 841
QY 798 E-----INATGHQOTLLVPSDEGATVLPPIRPTHGEIPIYVYALS-----PTA 841
DB 842 EQAPHCICANGRO-----TVSWAVPTPKSLGAVNFTVASALBSQELCGTEVPSV 890
QY 842 -----SDAVQMIIVKAGIEKYSQSILDLITDKRLOSTLKLSPSPRNTVTSERVQ 896
DB 891 PEHGRKDTVIKPLVEPGELEKETTFNSLCPSSGEVS---BELSLKLPENVEESASAS 947
QY 897 ITAIDVIGPSINGLASHLIMPYGCGEQNMIFANITYILDVLTQKKOLTDNLKESALF 956
DB 948 VSVIGLIDIGSAMONTNOLMPPYCGEQNMVIFANITYILDVLTQTLTPVSKALGY 1007
QY 957 MRQGYRELLYOREDGSPSAG-NDPS--GSTWLSAFVLRCELEADPYIDQNVLHRTY 1014
DB 1008 INTGYQOROLNMGHYGYSYTFEGRYGRNGNTWMTAFVLTQPAQARAVIFIDEAHITQAL 1067
QY 1015 TWLKHQSGNGBFMPGRVISHSELQGNKSPYTLTAYIVTSLGTRKQOPNIDVQESLHF 1074
DB 1068 IWLSPQKQKNGCFRSGSLNNAIKGVEDEVTLTASVITIALLEIPLTVTHVNVNNAFLC 1127
QY 1075 LES-----EFSRGISDNTYLLALITYALS VGS -PKAEALNMLTWRAEGBGQMFW--- 1124

```

```

DB 1128 LESAKTIAQEDDHG-SHYTVALALAYAPALAGNODKREKXVKSINEBAVKDNGSVHERP 1186
QY 1125 VSSSEKLSDSMOPR--SLDIEVAAYALSHLQO---TSRGIP-----IMRWLSNORSL 1175
DB 1187 OKPKAPVHFHPQAPSAVEKETSIVLA-VLTQAPATSDLSATNIVKMTIKQONAO 1245
QY 1176 GGFASDTDTVALKALSEFALMNTERTINQVTVGSPSPPLAVQ----- 1222
DB 1246 GGFSTQDTVALHALSKYGAATFL-RGKXAQVLTQSSGFFSKFQVDNNRLLQVYS 1304
QY 1223 ---PMAVNISANGFALCOLNVYTNKASGSSRRRSIONQAEFDLDAVK--ENKD 1275
DB 1305 LPELFGEXSMKVTGCGCYVLOTSLKYN-----LPEKEFPALQVOTLPQICD 1353
QY 1276 DLN-HVDLNVCTSPGCG---RSGALMEVNLISGFNWPSSAIDL--SEYKVEYDHG 1328
DB 1354 EPKATSFQISLSVSYTSSRSASNMAIVDVMVSGFIPLKPTVMKLEKSNHVSRTVESSN 1413
QY 1329 KLNLYLDSNETQFCVNIPIAVRNFVSNTOVASVIVDYEPBRAQVRSYNS 1380
DB 1414 HVLITLKVSNQTLISLFTVYQDVPRDLKRAIKVVDYETDEFALENYA 1465

```

```

RESULT 8
US-10-292-081A-12
; Sequence 12, Application US/10292081A
; Publication NO. US20030162202A1
; GENERAL INFORMATION:
; APPLICANT: Kenneth David Becker
; APPLICANT: Gonul Velicetlebi
; APPLICANT: Xin Wang
; APPLICANT: Randolph E. Tanzi
; APPLICANT: Lars Bertam
; APPLICANT: Aleister J. Saunders
; TITLE OF INVENTION: SINGLE NUCLEOTIDE POLYMORPHISMS AND MUTATIONS ON ALPHA-2-MACROGLO
; FILE REFERENCE: 37481-3323
; CURRENT APPLICATION NUMBER: US/10/292,081A
; PRIOR FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: 60/337434
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 1474
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-292-081A-12

```

```

Query Match 19.6%; Score 1440; DB 12; Length 1474;
Best Local Similarity 29.1%; Pred. No. 6.7e-109;
Matches 451; Conservative 266; Mismatches 551; Indels 284; Gaps 59;
QY 12 LVCVTAALAAVAPRFLVTPAGIIRPGANTIGVELLEHCPQVTVAKELTASNLTV 71
DB 15 LVLLPDPASVSGRQVWLVPSLIHT-ETTEKGCVLISYINETYVSALESVRGNSL 73
QY 72 -SVLEABG-VFEKSSFKTLTLPSPLSNADG---IYELRVGTQODELLFNSNTRLSFET 126
DB 74 FTDLAENDVLAHVAF-----AVKSSSNEVWFLVQVGPQOE---FKKRTVWVYN 124
QY 127 KRISVFIOTDKALYKPKQEVKFRIVTLFSDFKPYTSLNLI- IKOPKSNLIQOMLSQSD 185
DB 125 EDSLVEFVOTDKSIYIRGQTVKFRVVSMDENFHLNELLPLVYIOPDKNRIAQWOSFOLE 184
QY 186 LGVISTKFTOLSSHPLIDGMSIOVOND--QTYOSFOVSEVYLPRPEVTLQTPLYCSMNS 243
DB 185 GGLKQSPFLSSSEPPQSGYKVVQKSGGRTEH-PTVEEVLPRPEVQVTPKIIITILE 243
QY 244 KHLNGTITAKTYTKPKVKGDTVLTLPL-----SFMGK 276
DB 244 EEMNVSVGLTYTKPKVPGHVTYSICRKYSDASDCHGEDSOAFCEKFSQQLNSHCIFYQ 303

```

```

Qy 277 KKNITTKFKINGSANFSFNDEEMKNMDSNGLSBYLDLSSPGVEIIT-----TYESV 331
Db 304 VK--TKVFOUKR-----KEYEMK-----LHTEAQIOEGTVVELTGRSSEITRTI 347
Qy 332 TGISRNVTNVPFKOHDIIEFPDYTTVLKPSLNFTATVKTTRADGNQLTLEBRNNVVI 391
Db 348 TKLS-FKAVVDHFRQG---IPFFG-----QVRLVDGKGVPIP---NKVI 384
Qy 392 TTVGRNTEYWSGNSGNQKMEAVQ-KINYT--VPOSGTFKIEFP-----ILED 438
Db 385 FI-RGNEANY--SNATTDEHGLVQPSINTTNVMTGSLTVRVNKKDRSPCYQWVSEH 441
Qy 439 SELQKAYFLGSKSMVAHSLFKSPSKTYIOLKTRDENIKVGS---PFEIVSGNK--R 492
Db 442 EEAHTTAY-----LVFSPSKFVHLEPMHSHLPCGHQTQVQAHYILNGTLLG 489
Qy 493 LKELS--TMVVSROGLVAVG-----KONST---MPSLTENMTPKACIVVYIEDG 540
Db 490 LKKUSFYLLMAKGIYVTHGLLVKQEDMKHPSISIPKSDIAPARLLIYAVLPTG 549
Qy 541 EISDVLKIPYQVFNKNIKLYWSKVKAPESEKVSIRISVTQPSIYIGIVADKSVMLN 600
Db 550 DVIGSADYDVENCANKVDSLSPSSQSLPASHMLARTAA--PQVCALRAVDQSVLLMK 608
Qy 601 ASNDITMENVHEL-ELYNTGY-----YIGMFENS-----FA 631
Db 609 PDAELSSASVYVNLPEKDLTGFPGLPDQDEDCINRHNVYINGITYPVASTNEKOMYS 668
Qy 632 VFOEGGLVLTDAULTKXYIDGVYDMAE-----YARMEENEGHIVIHPSL 680
Db 669 FLEDMGLAFTNSKIRKPKMCPQLOQYEMHGPESLRVGFYSDDVGRCHARLVHEE-- 725
Qy 681 GSSPH--VRKHFPETWIMLDTNMGRYIYQEFEVTPDPSISWVATGVISEDGLT 737
Db 726 ---EHTEVTKYFPEWTIMDLVNVNAGVAVGTVPPTTEKAGAFCLSEDLGLTSS 782
Qy 738 TPVELQAFQPFIFELNLPYSVIRGEFALETTFNYLKDATEVYVILEKSDKFDILMTSS 797
Db 783 T-ASIRAPQPFVELTMYSVIRGEAFPLKATVNLPRKICRVSVQLEASAPFLAVPREK 841
Qy 798 E-----INATGHOQTLVPSDEGATVLEPIRPHLGEIPIVYALS-----PTA 841
Db 842 EQAPHICANGRQ-----TVSWAVTPKSLGNVNFVSAALESQELCTEVPVS 890
Qy 842 ---SDAVTOMILYKAGIEKYSOSITLDTNRLQSTLKTLSFSPPTVATGSEVQ 896
Db 891 PEHGRKDVYKPLVERPGLKEKETTNSLLCPGSEVS---EELSLKLPNVVESARAS 947
Qy 897 ITAIGDVLGSPINGLASLIRMPYGGGEQNMIFAPNIYILDYLTCKKQTLTNLEKALS 956
Db 948 VSVIGDILGSAHQNTQNLQMPYGGGEQNMVLPAPNIYVLDYLNBTQOLTREVSKAIGY 1007
Qy 957 MRQGYQRELLYQREDEGSFSAFG-NYDES-GSTWLSAFYLRCPLEADPYIDIDONVLRHTY 1014
Db 1008 LNTGQORQNTYKHVYSGYSTGERGRNQGTWMLTAFPLKTPFAQARAVIIFIDEAHITQAL 1067
Qy 1015 TWLKGHOKSNGEPMQGRVHISELQGNKSPVTLTAYIVTSLGVRKQPHIDVQESIH 1074
Db 1068 TWLSORQKQNCQFNSSGSLNNNAIKGVEDEVTLTASVITLALPLTPTTHPVVRNALPFC 1127
Qy 1075 LES-----EFSRGISDNYTLALITYALSVS--PKAKEALNMLTWAREGEGQWFM-- 1124
Db 1128 LESAMKTAQEGDGH-SHYTTKALLAYAFALAGNDKREVLKSLNEBAVKKDNVHMERP 1186
Qy 1125 VSSSKUSDSWQPR--SLDIEVAAYALISHFLQFO--TSEGIP--IKMWSLRQNSL 1175
Db 1187 QKPRAPVGHFEYEPQAPSALVEMTSTYVILA-YLTQAPATSEDLTSATVIVKMITKQNAQ 1245
Qy 1176 GGFASDTOTVATKALSSEFALMTERTNIVTVPSSPSPLAVVQ----- 1222
Db 1246 GGFSTDTOTVATLHLSKYGAATFL-RTKKAQVYTIOSSTGFSKFOVDNNRLLQOVS 1304
Qy 1223 ---PMAVNISANGFPAICQNTVNVYVYKAGSSRRRSRRIQNGEAFDLDAVK--ENKD 1275

```

```

Db 1305 LPELPEYSMKVYGEQGVYLOTLKKNYNI-----LPEKEEPPFALGVQTLPTQCD 1353
Qy 1276 DLN-HVDLANCTSESGRG---RSGMALMEVNLISGFVPEALS---SETVKAVEYDHG 1328
Db 1354 EPKAKHSFQISLSVSTYSGSSASNMALVDKMSGFLPKPYTKMLERSHNVSTEVSSN 1413
Qy 1329 KNLVYDSVNETQFCVNIPAVRNFKVSNTOODASVSIVDYVEPRROAVRSYNS 1380
Db 1414 HVLIVYDKVSNQTLSTLEFVYLOVVPVRDLKPAIVKYDYETDEPAIAEYNA 1465

RESULT 9
US-10-292-081A-13
; Sequence 13, Application US/10292081A
; Publication No. US20030162202A1
; GENERAL INFORMATION:
; APPLICANT: Kenneth David Becker
; APPLICANT: Gonul Velicelci
; APPLICANT: Xin Wang
; APPLICANT: Randolph E. Tanzi
; APPLICANT: Lars Bertram
; APPLICANT: Aleister J. Saunders
; TITLE OR INVENTION: SINGLE NUCLEOTIDE POLYMORPHISMS AND MUTATIONS ON ALPHA-2-MACROGLOB
; FILE REFERENCE: 37481-3323
; CURRENT APPLICATION NUMBER: US/10/292,081A
; CURRENT FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: 60/337434
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 1474
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-292-081A-13

Query Match 19.6%; Score 1440; DB 12; Length 1474;
Best Local Similarity 29.1%; Pred. No. 6,7e-109;
Matches 451; Conservative 266; Mismatches 551; Indels 284; Gaps 59;

Qy 12 LKCVTAALAVADGPRFLVTAAGIIRPGNVITGVLELHCPQVYKAEILKTASNLTV 71
Db 15 LVLVLPDASVSGKQPMVLPVSLHT-ETTEKGVLLSYLNTVYASLSVRGNRSL 73
Qy 72 -SVLEAEG-VFKGSEFETLTPSLPLNSADE---IYELRATGTODEILFNSNTRLSFET 126
Db 74 FTDLAENDVLHCVAF-----AVPKSSNEEVMFLTVQVGPPOE---FKKRTYVWVK 124
Qy 127 KRISVFIQTKALYKPKQEVKFRITVLPDFKPKYKTSNLTL-IKDPKSNLIQOMLSQOSD 185
Db 125 EDSLIVFQIDKSIYKKEQYKFRVVSMDENFHLNELIPVYIODEPKGNRIAQMOSEFOLE 184
Qy 186 LGVISTKTFQSLSPHILIGDWSIQOVND--QTYQSFQVSEYVLPKEVTLQTPLYCSMNS 243
Db 185 GGLKQSFPLSPSSPFGSVYVWVQKXSGGRTGH-PFTVEFVLPRKEVQTVTKIITILE 243
Qy 244 KHLNGITTAITYTKGYKGVDTYTLFPL-----SFMK 276
Db 244 EEMNVSVCGLYTGKVPVGHVTVSICRKYSDDASDCHGEDSQAFCEKFRSGQLNSHGCFYQ 303
Qy 277 KKNITTKFKINGSANFSFNDEEMKNMDSNGLSBYLDLSSPGVEIIT-----TYESV 331
Db 304 VK--TKVFOUKR-----KEYEMK-----LHTEAQIOEGTVVELTGRSSEITRTI 347
Qy 332 TGISRNVTNVPFKOHDIIEFPDYTTVLKPSLNFTATVKTTRADGNQLTLEBRNNVVI 391
Db 348 TKLS-FKAVVDHFRQG---IPFFG-----QVRLVDGKGVPIP---NKVI 384
Qy 392 TTVGRNTEYWSGNSGNQKMEAVQ-KINYT--VPOSGTFKIEFP-----ILED 438
Db 385 FI-RGNEANY--SNATTDEHGLVQPSINTTNVMTGSLTVRVNKKDRSPCYQWVSEH 441

```

```

QY 439 SFLQKAYFLGSSKSAVHSLFKSPSKTYIOUKTDENIKVGS-----PPELVVSGNK--R 492
Db 442 EBAHHHTAY-----LVFSPSKSVHLEPMSHSLPCCHTQTVQAVILINGTLIG 489
QY 493 LKELS--YVWVSRGOLVAVG-----KONST----MFLTPGNSMTPKACVIVVYIEDG 540
Db 490 LKKLSFYILIMAKGIVRTGTGLVKGQDMKGHSISIPVKSIDIAVPAKRLITVAVLPFG 549
QY 541 EILSDVLKIPVQVLFKNKIKLYMSKVKAPSEKVSLSVTPDPSIVGIVAVDKSVMLN 600
Db 550 DVGISAKVDENCLANKVDLSFSPQSIPASHAHLRVTA--PGVCALRAVDQSVILMK 608
QY 601 ASNDITMENVVHEL--ELVNTGY-----YLCMFENS-----FA 631
Db 609 PDAELSSASVYVNLPEKDLTFPGPLNDDEDICINRHVYINGLTYTPVPSSTNEKMDYS 668
QY 632 VPOEGCLWVLTDLNLTQYIDGVONAB-----YARFMEENGHI VDIHDP 680
Db 669 FLEDGKLAFTNSKIRKPMCPLOQOYEMHGPEGLRVGFESDVNGRHAUVHVE-- 725
QY 681 GSSPH---VRKHPPTWIMLDITNMGRIYQEFVTVPOSITSVATGFIYSEDLGLT 737
Db 726 ---PHTETVRKYPFETWIMDLVAVNASAGVAVTVPTDITMKAGAPCLSDAGLIGSS 782
QY 738 TPVELQARQPPFIPLNLPYSVIRGEBFALEITFNVLKDATEVKVYIEKSDKFDILMTSS 797
Db 783 T-ASIRARQPPFVELTMTPEVIRGEFTLKATVNLFPKICIVSVQLEASPFLAVPEK 841
QY 798 E-----INATGHQOQLVPSDEGATVLPPIRPTHGEIPITVTAIS-----PTA 841
Db 842 EDAHPICLNGRO-----TVSAVNTKSGNNAFTVSAELSESQELCTEVPBV 890
QY 842 ---SDAVTQMLVKAEGIEKYSQSILDLTDNRLOSTLKLTSFSPPTVTGSEK 896
Db 891 PEHKGDIYIKLVLVEPGELEKETPNSLCCSGGEVS---EELSIKAPPNVVESSARAS 947
QY 897 ITAIDVIGPSINGLASLIRMPYGCGEQNMNIPANVYILDYLTKKQJTDNLKEREALSF 956
Db 948 VSVLDIIGSAMQNTQNLQMPYGCGEQNMVLFAPNIVYLDVLTNETQUTPEVSKALGY 1007
QY 957 MROGTQRELLYOREGSPSAG--NTPDS--GSTWLSAFVLRCLLEADPYIDIDQNVLHRTY 1014
Db 1008 INTGYORQNLNKHVGYSTFSEYGRNOGNTWLAFLVLTAKPAQARAVIDEAIHTQAL 1067
QY 1015 TWLKCHOKSNGEFPMPGRVYIHSLELOGNKSPTLTAYITSLTGRKKQOPNIDVOESIH 1074
Db 1068 IWLQRQKONGCFRSGSSILNNAIKGVDEVTLSAYITIALLEIPLVTYTHVVRNALFC 1127
QY 1075 LES-----EFSRGISDNYTLLALITYALS VGS--PRAKEALNMLTWRAEOEGMQFW-- 1124
Db 1128 LESAKMTAQEGDHG--SHYTTKALLAVAFALAGNOKRKEVLKSLMBEAVKXONSVMHERP 1186
QY 1125 VSSSEKLSDSQOPR--SLDIEVAAYALSHPLQFO--TSEGI P---IMRLSQRNSL 1175
Db 1187 QKPKAPVGHFEPQAPSAEVEMTSYVLLA--YLTAPAPTSBDLTATNIVKMITRQONAO 1245
QY 1176 GGFASDTQPTVVALKLSERFALMNTERTNIQVTVGPSGSPPLAVVQ----- 1222
Db 1246 GGFSTQDPTVVALHLSKIGALFT--RTGKAQAVITIQSGTSSKRFQVDDNNRLLLOVS 1304
QY 1223 ---PMAVNISANGGFALQOLAVVYVNASGSSRRRSIQOEAFTDLDAVK-----ENKD 1275
Db 1305 LPELGEYSMKVLTGGCVYLTQTSKKNI-----LPEKEEPFALGVQTLPTQCD 1353
QY 1276 DLN--HYDLAVCTSFSGPG---RSGMALMEVNLISGFMVSEALISL---SETYKAYEYDHG 1328
Db 1354 EFKAPTSFQISLSVSTYGRSASNMAYDVKKVSGEIPLKPVTXKMLERSNHNSRTEVSN 1413
QY 1329 KLNLTDSVNETQFCVNI PAVNFKVSNTOQDASVSIQVDEYEPBRAQVNSYNS 1380
Db 1414 HVLITLDKVSQTLISLFTVLODVPVRDLKPAIVKVDYIETDERPALAEYNA 1465

```

```

RESULT 10
US-10-331-496A-38
; Sequence 38, Application US/10331496A
; Publication No. US20030228305A1
; GENERAL INFORMATION:
; APPLICANT: PRANTZ, GRETCHEN
; APPLICANT: HILLAN, KENNETH J.
; APPLICANT: PHILLIPS, HEIDI S.
; APPLICANT: POLAKIS, PAUL
; APPLICANT: SMITH, VICTORIA
; APPLICANT: SPENCER, SUSAN D.
; APPLICANT: WILLIAMS, P. MICKY
; APPLICANT: WU, THOMAS D.
; APPLICANT: ZHANG, ZEMIN
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; FILE OF INVENTION: TREATMENT OF TUMOR
; FILE REFERENCE: P5014R1-PCT
; CURRENT APPLICATION NUMBER: US/10/331,496A
; PRIOR FILING DATE: 2002-12-30
; PRIOR APPLICATION NUMBER: US 60/345,444
; PRIOR FILING DATE: 2002-01-02
; PRIOR APPLICATION NUMBER: US 60/351,885
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: US 60/360,066
; PRIOR FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: US 60/362,004
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/366,869
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: US 60/366,284
; PRIOR FILING DATE: 2002-03-21
; PRIOR APPLICATION NUMBER: US 60/368,679
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US 60/404,809
; PRIOR FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: US 60/405,645
; PRIOR FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 95
; SEQ ID NO 38
; LENGTH: 1474
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-331-496A-38

Query Match 19.6%; Score 1440; DB 12; Length 1474;
Query Local Similarity 29.1%; Pred. No. 6.7e-109;
Matches 451; Conservative 266; Mismatches 551; Indels 284; Gaps 59;

QY 12 LVCCTALAVAPGRFVTAAGIIRPGANTTIGVLELHCPQVTVAAELTKTASNLTV 71
Db 15 LVLVLPDASVSGRQVAVVLPSSLHT--ETTEKGCVLSTYANETVYASLESVRGRSL 73
QY 72 -SVLEAG--VEKSGFKTLPLSLPLNSADE--IYELRYTGRTODEILFNSSTRLSFET 126
Db 74 FTDLAENDVHCVAF-----AVKSSSNEVWELTVQVKGPTQE--FKKRTTVWKN 124
QY 127 KRISFIOTDALKAKPKQKQEVKRIYVLTSPDKPYKTSNLIL--IDPKSNLLOQLSOOSD 185
Db 125 EDSLVEFQVTDISIKYPGTIVFRVAVSDENHPLNELPLVYIDDPKGNRIAQOQSFQLE 184
QY 186 LGVISTKQFQSLSHPLDGMSIQOVND--QTYQSFQVSEYVLPRKEFTVLTQTPYCSMNS 243
Db 185 GGLQKQSFPLSSEPPQSGYKVVQKSGGRTEH--PFTVEEVLVLRKEFVQVVPKIIITILE 243
QY 244 KHLNGTITAKYTYGKPVKGVDTLTFPL-----SEWKG 276
Db 244 EEMWVSVGLTYTGKPVGHTVSIICRYSDASDCHGEDSOAFCEKFSGQLNSHGCFYQ 303
QY 277 KKNITKFKINGSANFSNDEEMKNWDSNGLSBEYDLSRPGVEILT-----TYTESV 331
Db 304 VK--TKVQFLR-----KEYEMK-----LHTEAQIOEGTVVELTGRQSEITRTT 347
QY 332 TGISRNASTNVFQKHDIYIEFPDYTTVLKPSLNFATVKTTRADGNQLTLEBRNNVVI 391

```

```

Db 348 TKLS-FVVDHFRG--IPFG-----QVRLVDGKVPF--NKVI 384
Qy 392 TVTQNTREYMSGNSGNGKMEAVQ-KINTY--VPOSTFIEEP-----LLEDS 438
Db 385 FT-RGNENANY--SNATTEHGLVQFSINTTNVMTSLTVAVNKKDSCPQYGVMBEER 441
Qy 439 SELQKAFGLGSKSMVAHSLFKSPSKTYIOLKTRDENIKYS--PFLVWSGNK--R 492
Db 442 BEAHHTAV-----LVFSPSKSFVHLEPMHSHLPCGHQYQYAHYILNGTLLG 489
Qy 493 LKEIS--YMVVSRQOLVAVG-----KONST--MESLTPESWTPKACVITYIEDDG 540
Db 490 LKKSIFYLIMAKGIVTGTGHLVVKQEDMKHFSISIPIKSDIAPARLLIYAVLPTG 549
Qy 541 EIIIDVAKIYQVLPKFKIKLYMSKVABSEKYSIRISVQPSIYQIVAVDKSVMNM 600
Db 550 DVIDSARYDVENCILANKVDLSFSPSOLPASHAHLRVTAA--PQVCALRAVDOSVLLMK 608
Qy 601 ASNDITMENVHEL-ELYNTGY-----YLGMPMS-----FA 631
Db 609 PDALSSSVYVNLPEKDLTGPGLANDQDDDCINRNVYINGITTPVSTNEXOMYS 668
Qy 632 VFQCGMLTDLANTKDYIDGYDNE-----YARMEENEGHIVIDHPSL 680
Db 669 FLEDMGLKAFNSKIRKPKCPOLQOYEMHPEGLRVGFESDYMGRAHVAHEE-- 725
Qy 681 GSSH--VRKHPEWTWIDTNNNGRYOFEFTVPDSTISWATGVSIEDGLT 737
Db 726 --PHETVRKYFPEWTWIDLVVNSAGVAVGTVPDITTEWKAFCISDGLGIS 782
Qy 738 TPVELQAFPEFIFLNPYSVIRGEEFALITIEFYLKDATEVYKITEKSKPILMTSS 797
Db 783 T-ASLRAPQPFVELTMYSVIRGEAFILKATVNLKPCIRVSVQLEASAFIAPVPEK 841
Qy 798 E-----INATCHOQTLVPSSEDAVLPPIRPHLGBIPITVYALS-----PTA 841
Db 842 EQAPHCICANGRO-----TVSWAVTPKSGVNFVSALESQELCTEVPVS 890
Qy 842 -----SDAVTOMILVKAIGIKYSOSLILDLTNRLOSTLKTISFSPPTVYTSERVQ 896
Db 891 PEHGRKDTIVIKPLVPEPGLKETTFNSLLCPSGEVS--BELSLKLPVNVESARAS 947
Qy 897 ITAIGDVLGSPINGLASLIRMPYCGGEONMIFANITYIDYLTKKKQLTNLEKALSIF 956
Db 948 VSVUGDILGSMQNTQNLQMPYCGGEONMVLPAFNITYIDYLTQQLTREVSKALGY 1007
Qy 957 MRQGYORELLVQREDSFSAFG-NYDPS-GSTWLSAFVRCFLBAPYIDIDQNVLRHY 1014
Db 1008 LNTGYQRLVYKHYDGSYSTGERYGRNQGTWLTAFVLTKEFAQARAVIFIDEAHITQAL 1067
Qy 1015 TWLKGHQKSNCFMDPRGVHISELOGKNSPYTLATYVLSLGRYKQPNIDQESLHF 1074
Db 1068 IMLSORQKDCNCFSSGSLNNNAIKGVEDEVTLISAVITLALBIPLVTHPVVRNALFC 1127
Qy 1075 LES-----EFSRGISDNYTLATITVALSSVGS--PKAKEALNMLTWABEGOGMFM-- 1124
Db 1128 LESAMKTAQEGDHC-SHYTALALAYAFALAGNODKKEVILKSLNEAVKXKONSVMHERP 1186
Qy 1125 VSESEKLSDSWQPR--SLDIEVAAYALSHFLQO--TSEGP-----IMKWSRQNSL 1175
Db 1187 QKPAAPGHFEPQAPSAEVMISYVILA-YLTQAPATSEDLASATNIVIMIRKQNAQ 1245
Qy 1176 GGFSTQDTTVALKALSFEFALMTERTNIOVITYTPGSSPEPLAVVQ----- 1222
Db 1246 GGFSTQDTTVALKALSFEFALMTERTNIOVITYTPGSSPEPLAVVQ----- 1222
Qy 1223 ---PMAVNISANGGFAICOLNVVYVNVKASGSSRRRSIQNOEAFDLDVAVK--ENKD 1275
Db 1305 LPELPGEYKMYTBEQCYIOTSLKTYI-----LPEKEEPFALGVQTLPOICD 1353
Qy 1276 DLN-HVDLNVCTSESGP--RSGMALMEVNLISGFMPVSEALSL--SETVKKVEYDHG 1328

```

```

Db 1354 EPRAHSPFQISLSVSYTGRSASNAIMAVDKVMSGFIPLKPTVKMLERSNHVSRTVESSN 1413
Qy 1329 KMLVYDLSVNETQFCVNIAPVRFKYSNTQDASVSIYDYEPERQAVRSYNS 1380
Db 1414 HVLITDKNSQTLISLFTYVLQDVPRDLKPAIVKYDIYETDEPAIAEYNA 1465

RESULT 11
US-09-756-247-4
: Sequence 4, Application US/09756247
: Publication NO. US20030180722A1
: GENERAL INFORMATION:
: APPLICANT: Godbole, Shubhada D
: APPLICANT: Boyle, Bryan J
: APPLICANT: Mize, Nancy K
: APPLICANT: Deng, Cenhua
: APPLICANT: Goodrich, Ryle
: APPLICANT: Arterburn, Matthew C
: APPLICANT: Zhou, Ping
: APPLICANT: Tang, Y. Tom
: APPLICANT: Liu, Chenghua
: APPLICANT: Yeung, George
: APPLICANT: Demarec, Radcoje T
: TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO ALPHA-2-MACROGLOBULIN-LIKE POI
: FILE REFERENCE: HYS-31CIP
: CURRENT APPLICATION NUMBER: US/09/756,247
: PRIOR FILING DATE: 2001-01-08
: PRIOR APPLICATION NUMBER: 09/649,167
: PRIOR FILING DATE: 2000-08-23
: PRIOR APPLICATION NUMBER: 09/540,217
: PRIOR FILING DATE: 2000-03-31
: PRIOR APPLICATION NUMBER: 09/684,711
: PRIOR FILING DATE: 2000-10-06
: PRIOR APPLICATION NUMBER: 09/560,875
: PRIOR FILING DATE: 2000-04-27
: PRIOR APPLICATION NUMBER: 09/496,914
: PRIOR FILING DATE: 2000-02-03
: NUMBER OF SEQ. ID NOS: 41
: SOFTWARE: PatentIn version 3.0
: SEQ. ID NO. 4
: LENGTH: 1508
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-09-756-247-4

Query Match 19.6%; Score 1437.5; DB 12; Length 1508;
Best Local Similarity 27.1%; Pred. No. 1,le-108;
Matches 422; Conservative 274; Mismatches 552; Indels 307; Gaps 45;

Qy 20 LAVAPG-----PRFLVTAQPIIR-PGQNVYIGVELLEHCPQSVTVABELKTASNLTVSV 73
Db 10 LALSPIAEELPYLVLTLPRLNLP-----SVQKVCULDSBGSDVAFYTLERKDTOKL 65
Qy 74 LEAGVEFKSFFTLTLPSPILNSADEIYELRYGRTODEILFNSSTRSFETKRSVFL 133
Db 66 LEVSGLKRRHLHICISFLVPPAGTEBVAITIRSG-VGNNISPEKKKVLIQQNGTFLV 124
Qy 134 QTDKALYKQKQVKNFVTLFSDPKFYKTSNLTL-IKDKSNLIQOMLSQOSDLGYISKT 192
Db 125 QTDKPLTTPQOQYFPIVITWDSNFFVNDKYSWVELQDPNSNRIAOMLEVPPOGIVDS 184
Qy 193 FOLSSHPIIGDWSIQVQNDQTYQSFQYSEVYLPRFEVTLQPLVCSNMNSKHLNGTITA 252
Db 185 FOLAPEAMIGTYI--VAVAEGKTFGTFSEVYLPKFKVAVPEKELSTVQESFLYKIC 242
Qy 253 KYTYGKPVKGDVTLTFLPSFMGKKKNTYTFKINGSANESFNDDEMKVMDSSNGLSSEY 312
Db 243 RYTYGKPMGAVQVSV--QKAN-----TWYREVEREQLPDKCRNLISGQ 285
Qy 313 LD-----LSSP-----GVEILTYTESVTGISRNVSINVFPKQHDYIIEFF 354
Db 286 TDKTGCSAPVDWATPDLIGAYSHQINIVATVVEGTVGEANATONIIYISPMGSMTFE 345

```



QY	355	YTTVTLAKSLNFTAVKTVTRADGNQDNLTEBRNNVY-----TYTQGN-----	397
Db	346	DTSNFYHNFPPSGKIRVRGHDSFL-----KNHVLVIYVINGTGNQTLTDNNGLAP	400
QY	398	---YTEWSSGNSGNQKMEAVOKINITYVPOSGTFKIE-----FPILEDSESLQKAY	446
Db	401	FTLETSGNNGTD-----VSLBCKRQMEDLVNPEQVPRYQNAYLHLPF	445
QY	447	FLGSKSSMAVHSLFKPSKTYIQLKTRDENIKVGSPEFLVY-----SGNKKLKLXS	497
Db	446	YSTSRFFLGIRHL-----NGPLKCGQPEQVILVDYIDPADASPDQSI-SFS	490
QY	498	YMWVSROGLVAVGKON-----STMFSLTPENSMTPPACVIVYIYEDDEILISVL	547
Db	491	YLLIGKSLVMEGQKHLNSKKKGLKASFSLSTLFTSR LAPPSLTVIYAFPSGGVVAOKI	550
QY	548	KIPQVLVFNKIKLYMSKYKABPSKSVLARI-SVQPDLSIVGIYVADKSVNLMNSDITM	607
Db	551	QPSYGMCTDNQVSLQFSPSQOLPGAEBELQLOAA-PSGLCALRAVDSEVLILRP-----	603
QY	608	ENVVHELELVNTGYVIGMF-----	626
Db	604	-----DRLSRSRYV-GMFPWGYHYPYQVAEYDQCPVSGPMDPFQPLIDMPQSHSOR	657
QY	627	-----MNSFAYQECGLMWLTDANTTKDYIDGVYDNAEYA-----ERFME	666
Db	658	SIWRPSPSEGTDLFSPFRDVGKLKLSNAKIKP-VDCSHRSPREXTAMGGCPHEAES	716
QY	667	ENEGBIYDHFSLGSSPHVKRHPETMIWLDTMNGYIYQEFVTVYDSTISWATGEV	726
Db	717	STPLHQE-----DSQVROYPEPTEWLMOLFPLGNSGKAHVHTVPDATTETWMAQFC	768
QY	727	ISEDIGLGLTTTPYELQAFQEPFIFLNPYSGVIRGESEALETTFNYLKADATEVYILEK	786
Db	769	TSQGRGFLSPT-VGLTAKRFPFDULTPYSGVIRGESRLTRATIFNYLKDCIRQYTDLAK	827
QY	787	SDKFDILMTSSEINATGHQOQTLVPSBEDGATVLPFIRPETHGEIPIYVYAL-----	837
Db	828	SHEYOLESWADS-----QTSSCLCADDAKTHHWNITAVKLGHNIFTSTKILDSNEPCG	881
QY	838	-----SPTASDAVTOMILVKABG--IEKYSOSILDLTDNRLOSTLKLTSFPPNT	888
Db	882	GQKGFVPQKGRSDPLIKFVLVAKPBGVLVEKTHS-----SLICPKKVASSESIVLELPYDI	936
QY	889	VTGSERVQITRIG-----DYLGPSINGLASLIRMGYGEQONMI	927
Db	937	VPDSTKAVIYVLGQOELTIDSEKRRKMEAAKVMRIKMTALQNDLGVLOMPSGGCEQNMV	996
QY	928	NFASNIIYLLDYLTKKQKQTLDMKEKALSFMRGQYORELLYOREDSPAFGNDPDSGSTM	987
Db	997	LFAIIIVLYLKEKAGLITEIRSAVGFLEIGYOKELMYGHSNSYAFPERDONGTWT	1055
QY	988	LSAFVLRCELEADPYIDIDQNLVLRHTYTWLKGHQKNSGEMPDGCRVHISELQGNKSPVT	1047
Db	1057	LTAEVTCFCGQAKFIFIDPKNIQDALKMMMGNOLPSCGYANVGMLHTAMKGVGDVDEVS	1114
QY	1048	LTAIYVLSLGRYKQYONIDVOESHFLFSEFSKDISNNTYLLALITYLSSVSGSKABG-	1106
Db	1117	LTAIVYALLTEMGXDVDPMVSGQIRCKNS-ATSTTLYYQALLAATYFSLAGEMDINNI	1175
QY	1107	ALANVLTRABOEGQMFVWSESSEKLSDS--W-OPRSLDIEVAAYALLSH-----FLOQOT	1158
Db	1176	LLKQLODQALISGESIYWSQKPTBSNMSPMSEPAUVDELTAVALYLLQTLKPSLTQKEI	1235
QY	1159	SEGIPIKRWLSRQNSLGGFASDQDTVALKALSEFA--ALMNTERTINIQVYVT-----	1210
Db	1236	AKATISIVAMLAKONAHYAGFSSTQDTVALQALAKYATTAAMPSEBEINLVKSTENFORT	1295
QY	1211	-GPPSPSLAVVQ-----PMAVINISANGFPAICQLVNVYNNVAKSSGSRKRSRIQONOA	1265
Db	1296	FNIGSVNVLVFOQDTLPNVPGMYTTLASGQGVVQYVATLRANILPP-----TNMKT	1344

```

QY 1264 FDLVDVAK---EKKDLDLNDVLDNVCSTFGCP-GRSGMAMLEVNLSGF--AMPSEALISL 1316
Db 1347 FSLSVELGKARCEPPTSPRLTLTIHTSYGSSSSSMALVEYKMLSGFSPMGCTNOULL 1406
QY 1317 SE-TVKVEYDHGKLANLYLDSVNETOPCVNIIPAVRNPKVSNTODASVSYDYEPP 1370
Db 1407 QQPLVKVKEVGDTPLNLTLYDELKKNQTYFTTISQSVLYTNLKPATIKVYDYLPL 1461

RESULT 12
US-09-756-247-23
; Sequence 23, Application US/09756247
; Publication No. US20030180722A1
; GENERAL INFORMATION:
; APPLICANT: Godbole, Shubhada D
; APPLICANT: Boyle, Bryan J
; APPLICANT: Mize, Nancy K
; APPLICANT: Deng, Cenhua
; APPLICANT: Goodrich, Kyle
; APPLICANT: Atteidurn, Matthew C
; APPLICANT: Zhou, Ping
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Yeung, George
; APPLICANT: Dmanac, Radjoje T
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO ALPHA-2-MACROGLOBULIN-LIKE PO
; FILE REFERENCE: HVS-31CIP
; CURRENT APPLICATION NUMBER: US/09/756,247
; PRIOR FILING DATE: 2001-01-08
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/684,711
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 23
; LENGTH: 1450
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-756-247-23

Query Match 19.5%; Score 1433; DB 12; Length 1450;
Best Local Similarity 29.0%; Pred. No. 2.5e-108;
Matches 450; Conservative 265; Mismatches 551; Indels 284; Gaps 59;

QY 12 LACCTAALAVAPRPRFLVAPGIRGCAVNTIGVELLHCPQVYKAEILKTASNLTV 71
Db 2 LTVALLPTDASVSGRQVWLVLPISLHT-ETTEKGCYLLSTLNERTVVASLESVGRGRSL 60
QY 72 -SVLEAG-VFEKGSFRTLPLPSLPLNSADE---IYELRTGRTODEILSNSTRLSFET 126
Db 61 FTDLAEANDVLHCVAF-----AVPKSSSNEEYMLFLTVQKSGTQE---FKRTTVAVNK 111
QY 127 KRISVFIQTDKALYKPKQEVKPRIVTLFSDPKDYKTSNLIL-1KDPKSNLIQWLSQOST 185
Db 112 EDSLIVFQGTOKSIYKPSQGTAKFRVVSMDENFHLNELILPLVYIODPKGNRIAQWOSFOLE 171
QY 186 LGVSKTFQSLSPHPIIGDWSIOYQVND--QTYVQSPQVSRXYLPKRFTVLTQTLPLXGSMNS 243
Db 172 GGLNQSFSPSSPEPQSGSYRVVQKSGGRTEN-PTVEEFVLVPRKEVQVTVPRKIIITILE 230
QY 244 KHLNGITTAITYTKYQKPKVGDVTLTFLPL-----SFWGK 276
Db 231 EEMNVSVCGIYTGKVPGRHVTVSICKRYSADSDCHGSDSQAFCEKFSGQGLNHSNGCFYQQ 290
QY 277 KKNITTFKINGSNANFSENDSEMKVNDSSNGLSEVYLSSPGPVEILLT-----TVTESV 331

```

```

Db 291 VK--TKVQLR-----KEYEMK-----LHTEAQIOEBGVTELKQSSBITRTTI 334
Qy 332 TGISRNVSTNPFKOHDIYIEFEDYTLVKSLNFTATVKTTRADGNOLTLEBRNNVYI 391
Db 335 TKLS-FVKVDSHFROG---IPFFG-----QVRLVDGKGVPIP---NKVI 371
Qy 392 TYTQNTYEWMSGSGNQGKEAVO-KINYT--VPQSTFKIEFP-----ILEDS 438
Db 372 FT-KENENANY--SNATDEHGLVOPSTINTNVKSTSLVRNRYKDRSPCYQWMSSEH 428
Qy 439 SELQKAYFLGSKSSMAVHSLFKSPSKTYIOLKTDENIKVGS---PPELVSGNK--R 492
Db 429 EEAHHTAT-----LVFSKSKPVHLPEPMSHLPCHGHTQTVQAHYIINCGTLLG 476
Qy 493 LKEIS--YMWVSRGOLVAVG-----KONST---MFSLTPENSWTPKACIYIYIEEDG 540
Db 477 LKLSFYLLIMAKGIVRTGTGHTGLVQEDMKGHSISIPVKSIDIAPVRLIYALVPTG 536
Qy 541 EIIISVUKIPQVLVKNKIKLYMSKVKAPEBSKSLRISVYQPSDVIQIVAVDSVNLN 600
Db 537 DVIGSAYXDVENCLANKVLDLSFSPQSLPASHAHLRYTAA--POSCVALRAVDQSVLLMK 595
Qy 601 ASNDITMENVHLEL-ELYNTGY-----YLGEMFNS-----FA 631
Db 596 PDAEISASSVYNLPEKOLTGFPFGLNQDDEDCINRNVYINGITTYPPVSTNEKMYS 655
Qy 632 VFQEGCLWVLTDAFLTXYIDGVYDAAE-----YARFMEENEGHIVDIDHESL 680
Db 656 FLEDGKLKAFNTRKIRKPRMCPQLOQYEMHGPEGLRAGVFESADVGRGHARLVHER--- 712
Qy 681 GSSPR---VRKGFPRTWIMLDTNMKRYRYOEELYVPPSISWATGAVISBDGLGTT 737
Db 713 ---PPTETVRKYFPETWIMDLVVNVSAGVAEYGVVPTPTITTEWKAGACLSBDAGLSS 769
Qy 738 TPVELOAOPFFIFINLPYSVYRGEPALETIFUYLDATEVKYIIEKSDKFDLMTSS 797
Db -770 T-ASLRAPQPFVELTMYPSYIRKGAFTLKATVNLVLEKCRIVSQLEASPAFLAVPEK 828
Qy 798 E-----INATGHQOITLVPSEDAVLPPIRPHLGEIPITYVALS-----PTA 841
Db 829 EQAPHICANGRQ-----TVSAVNPKSLGNVNFVTSABALSESOLCGTEVPSV 877
Qy 842 ----SDAVTOMILYKAGIEKSYQSILDLDTNRLOSTLKTLSFSFPNTVYGSSEVQ 896
Db 878 PEHGRKDVIPKILVPEPGLKETTFNSLCPSSGEVS---EELSILKPPNVVERASAS 934
Qy 897 ITAIGDVLPSPINGLASLIRMPYCGGEONMIFAPNIYILDTLTKKOLTNLKEKALSF 956
Db 935 VSVLADILIGSAMONTQNLQMPYCGGEONWLPFANNIYVLDYMETQOLTEVYSKALGY 994
Qy 957 MROGYORELLYQREDSFSAFG-NDPS-GSTWLSAFVLCFLEADPYIIDONVLRHTY 1014
Db 995 LNTGYQORLNTKHYDGSYSTGERGRNQGMWTLFAPVLKTFPAQORAVIIFIDEAHITQAL 1054
Qy 1015 TWLKGHQKSGEPMDPGRVHISELOGKNSPYTLTAIYVTSLLGTRYQPNIDVQESLHF 1074
Db 1055 IMLSQRKDNOCFRSSGSLNNNAIKGVEDEVTLSAYITILLLEIPLTVTHPPVANAFLC 1114
Qy 1075 LES-----EFSRGISDNYTLALITYALSSVGS-PKAEALNMLTWRAEOGCGNQF--- 1124
Db 1115 LESAMKTAQEBDHG-SHYTYKALLAYAPALAGNODKREVLKSJNEEAVKKNDSVHMRP 1173
Qy 1125 VSSSEKLSDSWQPR--SLDIEVAAYALSHFQFO---TSEGIP---INWLSRORSLSL 1175
Db 1174 OKPAPRVGHFEPQAPSAEVENTSVLLA-YITPAQATSDSLNSATNIYKMTIQOQNAQ 1232
Qy 1176 GGFASDTDTVALKALSEFPAALMNTERTNIQVTVTGPSSPAPLAVQ----- 1222
Db 1233 GGFSSSTQDTVALHLSKYGATFT-PRGKAAQVTIQSSGFSSKRYQVDDNNRLLQOVS 1291
Qy 1223 ----PMAVNISANGGFAICQNLNVYNNKASGSSRRKRSIONQAFDLDVAVK--EKKD 1275

```

```

Db      12292 LPELPGCEYMKATYGBGCYVLTQTSLKATNII-----LPEKBEFPFALGVQLPQCDD 1340
Qy      1276 DLA-HVDLVNCTGSPSPG---RSGMALMEYNLLSGFVNPBEAISL---SETVKKVEYDHG 1328
Db      1341 EKAHATSPQISLSVSTYSGRSASNNIAVDKWSGFLPLKPYTMMLEBSNHNVSXTEVSSN 1400
Qy      1329 KNLVYDSNENETQFCVNIAPVARNFKVSNMQDASVSIYDYEPERRQAVRSY 1378
Db      1401 HVLIIYIDKVSNOTLSLFTVYLDVPVBDLKPAIVKYIDYIETDEFALAEY 1450

RESULT 13
US-09-756-247-25
Sequence 25, Application US/09756247
Publication No. US20030180722A1
GENERAL INFORMATION:
APPLICANT: Godbole, Shubhada D
APPLICANT: Boyle, Bryan J
APPLICANT: Mize, Nancy K
APPLICANT: Deng, Genhua
APPLICANT: Goodrich, Ryle
APPLICANT: Aterburn, Matthew C
APPLICANT: Zhou, Ping
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenphua
APPLICANT: Yeung, George
APPLICANT: Dmanac, Radoje T
TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO ALPHA-2-MACROGLOBULIN-LIKE POLYMERIZATION
FILE REFERENCE: HYS-31CIP
CURRENT APPLICATION NUMBER: US/09/756,247
CURRENT FILING DATE: 2001-01-08
PRIOR APPLICATION NUMBER: 09/649,167
PRIOR FILING DATE: 2000-08-23
PRIOR APPLICATION NUMBER: 09/540,217
PRIOR FILING DATE: 2000-03-31
PRIOR APPLICATION NUMBER: 09/684,711
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: 09/560,875
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 09/496,914
PRIOR FILING DATE: 2000-02-03
NUMBER OF SEQ ID NOS: 41
SOFTWARE: PatentIn version 3.0
SEQ ID NO 25
LENGTH: 1491
TYPE: PR1
ORGANISM: homo sapiens
US-09-756-247-25

Query Match      19.5%; Score 1433; DB 12; Length 1491;
Best Local Similarity 27.1%; Pred. No. 2,6e-108;
Matches 419; Conservative 212; Mismatches 551; Indels 302; Gaps 44;

Qy      26 PFVLTABEIR-PCGNVITIGVELLEHCPBQVYVKAELKTANSLTYSVLEABEVFEGS 84
Db      4 PNYLVTLPRALNFP-----SVQKVCIDLSPGYSDVKFTVLTETKDTYKLLLEYSLKKRHL 59
Qy      85 FXTLTLPSLPLNSADEIYELAVTGRTODELIFNSNTRLSPEFKRISVFIQTKMLYKPKQ 144
Db      60 HCISLTVPPAGCTEVAITIRVSG-VGNNISPEKKKVLQROGNGTFVQTDKFLYTPGQ 118
Qy      145 EVKFRIVTLFSDPKPYKTSNLIL-1KDPKSNLIQWMLSQSDGLGVISKTFQLSSHPILGD 203
Db      119 QYVFRIVTMDSNFVPRNDKYSNVELQDPNSNRIAQWLELVPEQGIYDLSFLQAEAMLG 178
Qy      204 WSIOYQVANDQTYQSSQVSEYVLPKFEVTLQTLPLXGSMNSKHLNGTITAKTYGKPVKGD 263
Db      179 YT-VAANVAGKFTGTFSEVEYVLPKFEVEVEPEKELSTVQESFLVKICCRATYGKPMIGA 236
Qy      264 VTLTFLPLSFQWKKKNIITFTFKINGANSFNDSEMKKNVWDSNGLSLEYLD---LSSP- 318
Db      237 VQVAVSVC-----QKAN-----TYWYREVEVEQLDPDKRNLNSGQTDKTCFSAFV 279

```

QY 319 -----GPEVLLTVTESVTGSRNVSTNVFKOHDIIEFPDYTVLKPSLN 365  
DB 280 DMATFDLIGAYASHQINIVATVEEGTVEANATQVIYISPOGMSWTFEDTISFHPNRP 339  
QY 366 FRATYKVTADNOGLTEERRNNVY-----TVQRN-----YTEWSSG 405  
DB 340 FSGKLRVRGHDSFL-----KHLVPLVLYGTNGTFNQTLVTDNMLAPFLETISGMNGT 394  
QY 406 NSGNQKMEAVOKINTVTPSGTFKIE-----PFIEDSSLOKAYFLSKSSMAVH 457  
DB 395 D-----VSLBGFQMEDLYNPEQVPRYQNALHLRPPYSTTRSPFGH 439  
QY 458 SLFKSPSKTYLQKTRDENIKVSGPELVY-----SGNKRLELSYMWVSRGOLVA 508  
DB 440 RL-----NGPLKCGQPOEVLVDYIDPADASPOGEI-SFSYVILIGKSLVM 484  
QY 509 VGRON-----STMSLTPENSWTPKACVITYIIEDDGELISDLKIPQVLKPK 558  
DB 485 BQOKHNSKKKGLKASFSLSLFTSRLADPSPVLYIYAIPPSGGVADKIQFSVGMCFDQ 544  
QY 559 IKLYMSKVAEPSEKSLRISTVQPSIYGIYAVDSVNLMAANDITMENVHELKYN 618  
DB 545 VSLGSPSQOLPEAEVLOLOAA-PESLCALRAVDSEVILLRP-----DRELSN 592  
QY 619 TGYILGMF----- 626  
DB 593 RSVY-GMFPWYGHYQVAYEQCPVSGWDEPQPLIDPMQGHSSQSIIMRPSFSG 651  
QY 627 MNSFAVQECGLWLTDLANTDYIDGVYDNMEYA-----ERFMEENGHIVDHD 677  
DB 652 TDLFSFRRDVGKILSNNAKKR-VDCHSRSPESYAMGCGGHPAEFESTPLHQE--- 707  
QY 678 FSLGSSPHRKHPPEWIMLDNMGRIYQEPFVYTPDSITSVATGFIYSEDLGILTT 737  
DB 708 -----DSQVROYFPETWLDLPPIGSGKEAVHTVPDAITEKMAFSCTOSRGGLSP 762  
QY 738 TPEVLOAFQPFPLNLPSYVIRGEFEALETIFNLYKQATEKVIIEKSDKFDIMTSS 797  
DB 763 T-VGLTAFKRFVDLTLFYSVVRGESFRLATIFNLYKQCIKIVQDILAKSHQVLSWAD 821  
QY 798 EINATGHQOTLLVPSEDAVTLPFIRPTHGELIPITVTL-----SPYA 841  
DB 822 S-----QTSSCCADAKTHHWNITAVLGHINFTISKILDSNBPCCGQKFPVQKR 875  
QY 842 SPAYVOMILVKAEG--IEKSYSOSILDLTDNRLOSTLKTLSFSPPNVTYTSERVQITA 899  
DB 876 SDTLKRPVLVKEGVLVEKTHS-----SLCPKGVASVSLELPVDIVPSTKAYVTV 930  
QY 900 IG-----DVLGSPSINGLASLIRMPYGGGEONMIPAPNIYILDY 938  
DB 931 LKQLEIIDSEKRRMEAAKWMDINGTLLQNLGVLQMPSGCGEONMVLFAPIIYVLY 990  
QY 939 LTRKQLOTLNLEKALSFMRQGYORELLYQREDSFSAFANYDPGSGTWLSAFVLRCFLE 998  
DB 991 LERAGLLTEIRSRRAVGFLEIGYQKELAMKHSNGSYSAFGBDNGENTWTLFAVTKCFQ 1050  
QY 999 ADPIYDIDONVLAHTYTLKGHOKSNGEWFDEGRVHSELQGSNSPVTLLTAIYTSLIG 1058  
DB 1051 AOKFIPIIDKXNIDALKMAAGNOLPBGCVANVGNLHTAMKGVDEVSILTAVVTALLIE 1110  
QY 1059 YARKYONIVOSIHFLESEFSGISDNTYLLALITVALSVSPRAKE-ALMMLTMRBAQ 1117  
DB 1111 MGRDNDPWPVSOGLRCLKNS-ATSTTNLYTQALAYIFSLAGBMDIRNILLQOLDOQAIT 1169  
QY 1118 EGGMQFWSSSEKLSDS--W-QPSLIDIEVAAYALISH-----FLOFQTSSEGIPIMRWLS 1169  
DB 1170 SGGSIYWSQKPTPSSNAPSMESEPAAYDVELTAVALLAQTLKPSLQKEIATKNTSIYAWLA 1229  
QY 1170 RQNSISGASASTDTTVALKALSEFA--ALAMTERINIOVTV-----GSSSPSEPLAV 1220  
DB 1230 KQHNAGGSSSTODIVVALQALAKATYATYAMSEEBINLVKSTENPQRTFNIOQSVRLVFE 1289

QY 1221 VQ-----PMANISANGFPAICQNLVNVYKASGSSRRRRSIONDEAFDLDAVK--- 1271  
DB 1280 QODTLPNVPMYTLBASQGCYVQTVLRNVLPP-----TMMTFSLSVIGGAR 1340  
QY 1272 -ENKDDLHVDLNVCTSSSGP-GRSGMALMEVNLISGF--MVPSSAISLSE-TYKKVFEYD 1326  
DB 1341 CEQPTSPSSLTLTIHTSYVSGRSSSMMAIVEKMLISGSPMEGTQOLLQOPLYKKEVFE 1400  
QY 1327 HGKMLLYDSVNEQFCVNIPAVNFKVSNTOGASVSLVDYEP 1370  
DB 1401 TDTLNTIYDELIKNTQTYTFTISQSVLVNMLKPAITIKVYDIYP 1444  
RESULT 14  
US-10-292-081A-9  
Sequence 9, Application US/10292081A  
Publication No. US20030162202A1  
GENERAL INFORMATION:  
APPLICANT: Kenneth David Becker  
APPLICANT: Gonul Velicelbeli  
APPLICANT: Xin Wang  
APPLICANT: Randolph E. Tanzi  
APPLICANT: Lars Bertram  
APPLICANT: Aleister J. Saunders  
TITLE OF INVENTION: SINGLE NUCLEOTIDE POLYMORPHISMS AND MUTATIONS ON ALPHA-2-MACROGLO  
FILE REFERENCE: 37481-3323  
CURRENT APPLICATION NUMBER: US/10/292, 081A  
CURRENT FILING DATE: 2002-11-08  
PRIOR APPLICATION NUMBER: 60/337434  
PRIOR FILING DATE: 2001-11-09  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 9  
LENGTH: 1500  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-292-081A-9  
Query Match 19.4%; Score 1428; DB 12; Length 1500;  
Best Local Similarity 28.7%; Pred. No. 6.7e-108;  
Matches 458; Conservative 270; Mismatches 567; Indels 300; Gaps 62;  
QY 12 LVCCTALAAVAPGRFVTAAPGIRPGANVTIGVLELHGCQVTVKAEILKTASNLTV 71  
DB 15 LVLVLPDASVSGKQYVNLVPSLIHT-ETTEKGVLLSYNLETYVVASLESVAGNSL 73  
QY 72 -SVLEAEG-VFEKGSFKTLTLPSPLENSADE---IYELRVGTQODEILFNSSTRLSSET 126  
DB 74 FTDEAENDVLAHCAF-----AVPKSSNEEVMFLTVQVGPQGE--FKKRTTVWYKN 124  
QY 127 KRISVFIQDRAKYKQGEVFRIVTLFSDRKPYKTSNLIL-IDPKSNLLOQLSOQSD 185  
DB 125 BDSLVTQDTSIYKPGQTVKFRVMSMDENHPINELIPLVYIDPKNRIAQWQSFQLE 184  
QY 186 LGVISTKQLSHPILIGWMSIOVQVND--QTYYGOSFOVSEVVLPRFEVTLQTPLYCSMNS 243  
DB 185 GGLKQFSPLSSEPPQGSYKVVQKKSGRTEH-PFTVEEFVLKPEFQVTVPKITITILE 243  
QY 244 KHLNGTITAKYTYGKPVKGVDTLTFPLPL-----SFWGK 276  
DB 244 EEMNVSVCGLLTYGKRPVGHVTVSICRKYSDASDCHGBDSQAFCEKFSQQLNSHGCIFYQ 303  
QY 277 KKNITKTRKINGSANFSNDEEMKXWDSNGLSLEYLDLSSPGVEIIT-----TVTESV 331  
DB 304 VK--TKVFOLRK-----KEYEMK-----LHTEAQIOEGFVLELTVGRQSSSEITRTI 347  
QY 332 TGISNVSTNVFVKOHDIIEFPDYTVLKPSLNTATVAKVTRADGNQGLTEERRNNVY 391  
DB 348 TKLS-FKVDGHHFQGG--IPFG-----QVRLVDGKVPILP--NKVI 384  
QY 392 TVTQRNTEYWSGNSGNQKMEAVQ-KINTY--VPOSGTFKIEPP-----ILBDS 438  
DB 385 FL-RGNEANY--SWATTDEHGLVQFSINTTNWGTSLTVAVNVKDRSPCYGQWVSEEH 441

```

QY 439 SEOLKAYFLGSKSMVAHSLFKSESKTYIOLKTRDENIKVGS-----PFELVSGNK--R 492
D 442 EEAHTAY-----LVFSPSKSFVHLEBPSHSLPCGHQTQVQAHYILNGTILG 489
QY 493 LKETS---YMWVSROGLVAVG-----KONST-----MBSLTRENSMTPKACIVVYIEDG 540
D 490 LKLSFYLLIMAKGIVRTGTHGLLVKQEDMKHPSISIPKSIDAPARLLIYAVLPTG 549
QY 541 EIIDVLKIPVOLVFNKIKLYMSKVKAPESEKYSLRISYTOPDSIVIGIVAVDSKNIM- 599
D 550 DVIGDSAKYDVENCILANKVDSLSPSPSGSLPASHAHLKRYTAA-POSVCLRAVDSDVLML 608
QY 600 -----NAS-----NDITMENVHLELYNTGY-YLGMFMS-----FA 631
D 609 PDAELSSASYVNLPEKDLTGFPGLNDQDNEDCINRHNVIYINGITLYPVSTNEKDMYS 668
QY 632 VFQECGLMVLTDANLTKYIDGVYDNAL-----YAREMEHEGHIYDIHPSL 680
D 669 FLEBMGLKAFNNSKIRKPCPOLQOQYEMHGPBGLRVGFYSDDVGRGHARLVHEE--- 725
QY 681 GSPH---YRKHPETWIMLDTNMGYRYLOEFVTVPPDSITSWATGVISEDGLGLT 737
D 726 ---PHTEVTKYFETWIMDLVYVNSAGVAVGTVPPTTEWKAQAFCLSEDRGLISS 782
QY 738 TPVELQAFQPFIFLNFYSYIRGEFPALBITFNYLKDATERVYIIEKSDKPDILMTSS 797
D 783 T-ASLRARQPFVELTWPYSYIRGEAPFLKATVNLVYKICIRVSGVQLEASAPFLAVPEK 841
QY 798 E-----INATCHOQTLVPSDEGATVLPFIPHILGELPIVTVALS-----PTA 841
D 842 EDAHPCICANRO-----TVSWAVTPKSLIGNVNFVSALESQELCSTEVPSV 890
QY 842 -----SDAVTOMIIVKAGIEKYSOSILDLTONRLOSTLKTSPSPPTVYSGEVQ 896
D 891 PEHGRKDVIVKPLVEBGELEKETTFNSLCPSGEVS---EELSLKIPNVVESARAS 947
QY 897 ITAIGDVLGSPSINGLASLIRMPYGGGEONMINFAPNIYILDYLTKKQLTDNLKALSF 956
D 948 VSVLGDILGSAMQNTQNLQMPYGGGEONVLPAPNIYILDYLTNETQQLTEIRSKAIGY 1007
QY 957 MRQGYRELLYQREDEGSFASG-NYDS-GSTWLSAPFLRCFLLEADPIIDQVNLHRTY 1014
D 1008 LNTYQROQLNXYKHDGYSYTFGEERYGRQGTWTLAFPLKFAQARAYIFIDEAHITQAL 1067
QY 1015 TWLKGHOSNGEFPDQGVHSELOGGKSPVTLTAYVTSLLGYRKQOPIDQESLHF 1074
D 1068 IWLQORQKONGCFRSGSLNNALIKGVEDVTLISAYITTLALPLTVTHPVVRNALFC 1127
QY 1075 LES-----EFSRGISDNYTLALITYALSVS-GPKAKELNMLTWRADEGGMQFM--- 1124
D 1128 LESAMKTAQEGDHG-SHYTTKALLAYAPALAGNDKREVLKSLNBEAVKKNDSVHMERP 1186
QY 1125 VSSSKSIDSQWR--SLDIEVAAYALISHLQOQ--TSEGIP---IMWLSRQNSL 1175
D 1187 QKPAPOHFEYEPAPSAVEVEMTSYVLLA-YLTQAPAPTSBDLTSATIVIMITKQONAQ 1245
QY 1176 GGFASDTOTVALKALSEFALMNTERTINQVTVTGPSSPPLAVVQ----- 1222
D 1246 GGFSTQDTVALHALSKYGAATFT-RTGKAQVYTIQSGTFSSKFOVDNNRLLQOVS 1304
QY 1223 -----PMAVNISANGFPAICQNLNVVYVKAAGSSRRRSRISIOQEAFLDVAVK--ENKD 1275
D 1305 LPELPEGYSMTYBEGCVYLOTSIKYNI-----LEKKEFPALQVQLLPQTC 1353
QY 1276 DLN-HVDLNVCTSFSGPG--RSGMALMEVNLISGFVWPSAISL---SETVKVEYDHG 1338
D 1354 EPKAKHTFQIQLSYTSYTSRGSASNMAYIDVVMVSGFIPLKFTVQALERSNHVSRTVSSN 1413
QY 1329 KLANLYLSVNETQCVNIPARANKVSTODASVYVYPRQAVRSNYSVKLSSCD 1388
D 1414 HVLILDKVSNQTLSEFTVLQDVVPRDLKPAIVKVDYB-----TGLQQLSLSTM 1464

```

```

QY 1389 LCS-----DVQCRPCBEDGASGHHHSVITF 1416
D 1465 LLAAKILEMLEDHKAEXCFAGVLFSELHRRHVFLY 1499

RESULT 15
US-09-981-151A-10
: Sequence 10, Application US/09981151A
: Publication No. US20030212256A1
: GENERAL INFORMATION:
: APPLICANT: Edinger, Shlomit R
: APPLICANT: Gerlach, Valerie
: APPLICANT: MacDougall, John R
: APPLICANT: Malyankar, Muriel M
: APPLICANT: Smithson, Glenda
: APPLICANT: Miller, Isabelle
: APPLICANT: Peyman, John A
: APPLICANT: Stone, David J
: APPLICANT: Gunther, Erik
: APPLICANT: Ellerman, Karen
: APPLICANT: Shimkets, Richard A
: APPLICANT: Padigaru, Muralidhara
: APPLICANT: Guo, Xiaojia
: APPLICANT: Paturajan, Meera
: APPLICANT: Taupier Jr, Raymond J
: APPLICANT: Burgess, Catherine E
: APPLICANT: Zerhusen, Bryan D
: APPLICANT: Kekuda, Ramesh
: APPLICANT: Szytek, Kimberly A
: APPLICANT: Gangoli, Bha A
: APPLICANT: Fernandes, Elma R
: APPLICANT: Gorman, Linda
: TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
: FILE REFERENCE: 21402-168
: CURRENT APPLICATION NUMBER: US/09/981,151A
: CURRENT FILING DATE: 2001-10-16
: PRIOR APPLICATION NUMBER: 60/241,040
: PRIOR FILING DATE: 2000-10-17
: PRIOR APPLICATION NUMBER: 60/241,058
: PRIOR FILING DATE: 2000-10-17
: PRIOR APPLICATION NUMBER: 60/241,063
: PRIOR FILING DATE: 2000-10-17
: PRIOR APPLICATION NUMBER: 60/241,243
: PRIOR FILING DATE: 2000-10-17
: PRIOR APPLICATION NUMBER: 60/242,152
: PRIOR FILING DATE: 2000-10-20
: PRIOR APPLICATION NUMBER: 60/242,482
: PRIOR FILING DATE: 2000-10-23
: PRIOR APPLICATION NUMBER: 60/242,611
: PRIOR FILING DATE: 2000-10-23
: PRIOR APPLICATION NUMBER: 60/242,612
: PRIOR FILING DATE: 2000-10-23
: PRIOR APPLICATION NUMBER: 60/242,880
: PRIOR FILING DATE: 2000-10-24
: PRIOR APPLICATION NUMBER: 60/242,881
: PRIOR FILING DATE: 2000-10-24
: Remaining Prior Application data removed - See File Wrapper or PALM.
: NUMBER OF SEQ ID NOS: 160
: SOFTWARE: Patentln Ver. 2.1
: SEQ ID NO 10
: LENGTH: 1492
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-09-981-151A-10

Query Match 19.3%; Score 1421.5; DB 12; Length 1492;
Best Local Similarity 27.2%; Pred. No. 2.3e-107;
Matches 419; Conservative 269; Mismatches 559; Gaps 44;

QY 20 LAYAPG-----PFLYTABPILR-PCGNVYIGVLEHCGSYTVKAEILKTSNLTVS 73
D 10 LALSPALIELPVLVTLPARLNF-----SVQKCVLDLSPGYSVDKFTVLTETKDKTQKL 65

```

```

QY 74 LEAEVFEKGFETLTLPSLPLNSADEIYELVGTGTODEILFNSNLSFETKISVFI 133
Db 66 LEVSGJKKHLHCISLFLVPPAGGTEEVATIRVSG-VGNNISFEKKEKTLIRQGGTFV 124
QY 134 QTDKALYKROEKYKPIVTLFSPDFKYKTSINIL-IKDPKSNLIQOMLSQSGDLGIVIST 192
Db 125 QTDKPLVYTGQOYFRIVTMDSNFVVDNKTSMVELQDPNSNRILQWLEVBEOGIVDS 184
QY 193 POLSSHPIIGDWSIQOVNDQTYOSFOVSEY-----VLKPEFVTLQTPLYCSM 241
Db 185 POLAPEAMLGTYT--VAVAEKGTFGFSVEEYVLSFPLLLSSVLPKFEVVEVEPELEST 242
QY 242 NSKHLNGITTAKTYYTKGVKGVTLTFLPLSFGKKKNI TKFKJNGSANSFNDDEKMN 301
Db 243 VQESPLVKICCRYYTKGPMIGAVOVSVQ--QKAN-----TYWVREVEREQ 285
QY 302 WMDSSNGLSLEYLD--LSSP-----GPVELLTYYTESVTGIRSNSTNIF 343
Db 286 LDPKCRNLSGQTDKTCGFSAPVDMATEDLIGYAVSHQINI VAVVEEGFVEAMNATQNIY 345
QY 344 FKQHDYIIEFPDYTVLVKPSLNFATVKTBRADGNOLTEBRRNNVI-----TV 393
Db 346 ISPOGSMWFEDTNSYHNPFPSGKMLKFPQGVLPCKNHLVFLVITGTGTNGTQIV 405
QY 394 TORN-----YTERWSSGNSNGQKAEAVOKINVTVPQGTFKIE-----FPILED 438
Db 406 TDNNGIAPFTELETSGWNGTD-----VSLGKFGOMEDLVNPROVPRYYGN 450
QY 439 SELQKAYFLGSKSSNAVSLFKSPKTYIQLKTRDENIKVSSPFVLV-----SG 489
Db 451 AYLHLRPFYSTRSFGLHRL-----NGPLKCGQPOQVLDVYIDPADASP 496
QY 490 NKRLKELSYMVVRGQLVAVGKON-----STMFLTPENSMTPKACVIVYYIEDD 539
Db 497 DGEI-SFSTYLLGKSGLVMEGQKHLNKKGLKASFSLSLTFSRLAPPSLVITAFPS 555
QY 540 GEIISDVLCIPVQVLFKFKIKLYWSKVKAPESEKVSIRISVTQPSIVGIVAVDKSVIM 599
Db 556 GGVVADKIQFSVMCFDN-----QQLPGAVELQLOAA-PSGLCALRAVDESULL 605
QY 600 NASNDITMENNVHELELYNTGYLGMF----- 626
Db 606 RP-----DRELSNRSVY-GMFPWGHYPYQVAEYDQCPVSGPMDEPQPLIDPM 653
QY 627 -----MNSPAVROEGMLWLTQDANLTQYIDGVNDAEYA----- 661
Db 654 POGHSSQRSIIMRPSSEGTDLPSFRDGLKILSNAKIKKP-VDCSHRSPSTAMGGG 712
QY 662 ---ERFMENECHIVDIDHPSLGSSPHVKHFPETWIMLDTNMGYRIYQEFVTVPDSIT 718
Db 713 GHPEAFESSTPLHQAE-----DSQVHQYFETWIMDLFPIGNSGKEAVHVTPDALI 764
QY 719 SWVATGCVI SEDLGGLTTTVELQAFQFFIFLNPYSVINGEPALEITTFNYIKDAT 778
Db 765 EKKAMSFCTSOGRGFLSPT-VGLTAFKPPFDLTLPSVVRGSESPRLTATIFNYIKCI 823
QY 779 EYKVIIKSKDFDILMTSEINATGHOQLVAPSEBGAFLVLPINRTHLGEIPITYAL- 837
Db 824 RYQTDILANSHEYQLESWADS-----QTSCLCADDAKTHMNITAVKLGHNFTISTXI 877
QY 838 -----SPTASDAVTQMLLVKAGIEKSYQSQILDLTDNRLOSTLKTLSF 882
Db 878 LPSNEPCGQKGFVPQKGRSDTLIKRVLKPEGVLEKTHSSLCCKGKVAS--ESVSL 935
QY 883 SFPENTVTSERVQITAIQDVLAGPSINGLASLIRMPYCGCEQNMIFADNIYILDYTKK 942
Db 936 ELPPVDIVPSTKAYVTVLQDIMGTAQLNDGLVOMPSGCEQNMVIFAIIYVLQYLEKA 995
QY 943 KQUTDLKXKALSPMGOYORELHYOREDSFSAFENVDPGSGTWSAFVLCFLEADPY 1002
Db 996 GLITEIRIRAVGFLTIGYQKELMYHSGSYSAFERDNGNATWLTAFVTKCFQQAQKF 1055
QY 1003 IDIDQNVLHRTYTWLKGHQKNGEFPWPGRVIHSELQGNKSPVLTAYIVTSLGGRKY 1062

```

```

Db 1056 IFIDPKNIQDALKMWAGQULPSGGCVANVGNILHTAMKGVVDDEVSLTAYVTAALLEMGKD 1115
QY 1063 QPNIDVQSIHFLSEFSGRISDNVTLTALITYALSSVSGSPRAKE-ALNMLTWRAQEGCM 1121
Db 1116 VDDPVSQGLCKLQNS-RTSTNLTQMLLAVITFSLAGEMDIRNILLQLOQALISSES 1174
QY 1122 QFWVSSBSKLSDS--W-QPRSLDIEVAAYALISH-----PQOFQTSBGIPIMRWLSROPN 1173
Db 1175 IYWSGKPTPSSNASPMSRPAVDVELTAVALLAQUTKPSLQKELAKNTSIYAMLAKQHN 1234
QY 1174 SLGCFASIQDTTVALKALSEFA--ALNMTERTNIQVTVT-----GSSSPPLAVQ-- 1222
Db 1235 AYGGFSTQDVTVALQALAKYATTAAMPSEBINLVKSTENFQRTFNIOQVNRLVFOQDT 1294
QY 1223 ----PMAVNISANGGFALCOLNVVYVNVKASGSSRRRSIONQEAFLDVAVK---ENK 1274
Db 1295 LPNVGMYTTLSEGGQCVYQVTLKRYNLLP-----TNMKTFSLSVEIGKARCQOP 1345
QY 1275 DDLNHDVNLNCTSFSGP-GRSGMALMEVNLISGF--WVPSBAISSE-TYKKVEYDHGKL 1330
Db 1346 TSPRSLTLTHTSYVSGRSSSNNMAIVEYKMLSGSPMEGCTQOLLQCPVKKVEGCTOL 1405
QY 1331 NIYUDSVNETQPCVNI PAVRNFKVSNQDASVSIYDYEYEP 1370
Db 1406 NIYDELIKNTQVTFYFTISQSVLVTNLKPAITIKYDYLLP 1445

```

Search completed: January 15, 2004, 18:19:33  
Job time : 51 secs

**THIS PAGE BLANK (uspto)**



```

Matches 303; Conservative 0; Mismatches 295; Indels 6; Gaps 1;
QY 2649 ATTTCCTCTTAATACAGTACGAGTGGAGGAAAGTTCAATCACTGCAATTTGAGATGT 2708
Db 2845 ACTGCCACCAAAATGTGTAGAAATCTGCCAGCTTCTCTCAAGTTTGGAGACAT 2904
QY 2709 TCTTGCTCTTCATCAATGAGCTTAGCTCAATGATTCGATGAGCTTATGAGCTGTGTGA 2768
Db 2905 ATTAGGCTCTCCATGCAAAACACAAATCTTCTCAGATGCCATATGCTGTGGAGA 2964
QY 2769 ACAGAACATGATAATTTTGTCTCAATATTTAATTTGATTTATCTGACTAAAGAA 2828
Db 2965 GCAGAAATATGCTCTTGTCTCTTAATCTATGATCTGATTAATGAAACACA 3024
QY 2829 ACAACTGACAGATATTTGAAAGAAAGCTCTTTCATTTATGAGGCAAGTTACAGAG 2888
Db 3025 GCAGCTTACTCAAGATGCAAGTCCAGGCAATGGCTATCTCAACACTGTTACAGAG 3084
QY 2889 AGAATCTCTATCAGAGGAGATGCTCTTTCAGTGCCTTTGGG-----AATTATGA 2942
Db 3085 ACAGTTGACATCAAAACACTATGATGGCTCTACAGACCTTTGGGGAGCGATATGGAG 3144
QY 2943 CCTTCTGGAGCACTTGTGTGACGTTTGTTTAAGATGTTTCTTGAAGCCGATCC 3002
Db 3145 GAACGAGGCAACCTGGCTCACAGCTTGTCTGAAAGCTTTGGCCCAAGCTCGAGC 3204
QY 3003 TTACATGATATTTGATCAGATGATGTTTACAGAAACATACACTTGGCTTAAAGACATCA 3062
Db 3205 CTACATCTTCAATGATGAGACACATTTACCAAGCCCTCATATGAGCTCTCCAGAGCA 3264
QY 3063 GAAATCCACCGTGAATTTTGGATTCAGAGAGTGAATTCATGATGAGCTTCAAGGTGG 3122
Db 3265 GAAGGACAAATGGCTGTTTCAAGAGCTGGGTCACTGCTCAACATGCAATTAAGGGAGG 3324
QY 3123 CAATTAAGTCCAGTACACTTTACAGCTTATTTGTAATTTCTCTCTGGGATATGAAA 3182
Db 3325 AGTAGAAGATGAAGTGAACCTCTCTCCGCTATATACACATCCCTTCTGGAGATTCCTCT 3384
QY 3183 GTATCAGCTTAACATTTGATGCAAGAGTATCATTTTGGAGTGTGAATTCAGTAG 3242
Db 3385 CACAGTACTCACTGCTGTGTGCGCAATGCCCTGTTTGGCTGGAGTACGCTGGAAGAC 3444
QY 3243 AGGA 3246
Db 3445 AGCA 3448

RESULT 2
US-09-241-606-1
; Sequence 1, Application US/09241606
; Patent No. 6472140
; GENERAL INFORMATION:
; APPLICANT: Tanzi, Rudolph E.
; APPLICANT: Kovacs, Dora
; APPLICANT: Saunders, Aleister J.
; TITLE OF INVENTION: Alpha-2-macroglobulin Therapies and Drug Screening Methods for
; TITLE OF INVENTION: Alzheimer's Disease
; FILE REFERENCE: 0609.4460003
; CURRENT APPLICATION NUMBER: US/09/241.606
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4577
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: sig peptide
; LOCATION: (44)..(112)
; NAME/KEY: CDS
; LOCATION: (44)..(468)
; FEATURE:

```

```

; NAME/KEY: mat peptide
; LOCATION: (113)..(4468)
US-09-241-606-1
Query Match 2.6%; Score 110; DB 4; Length 4577;
Best Local Similarity 50.2%; Pred. No. 3.3e-21;
Matches 303; Conservative 0; Mismatches 295; Indels 6; Gaps 1;
QY 2649 ATTTCCTCTTAATACAGTACGAGTGGAGGAAAGTTCAATCACTGCAATTTGAGATGT 2708
Db 2845 ACTGCCACCAAAATGTGTAGAAATCTGCCAGCTTCTCTCAAGTTTGGAGACAT 2904
QY 2709 TCTTGCTCTTCATCAATGAGCTTAGCTCAATGATTCGATGAGCTTATGAGCTGTGTGA 2768
Db 2905 ATTAGGCTCTCCATGCAAAACACAAATCTTCTCAGATGCCATATGCTGTGGAGA 2964
QY 2769 ACAGAACATGATAATTTTGTCTCAATATTTAATTTGATTTATCTGACTAAAGAA 2828
Db 2965 GCAGAAATATGCTCTTGTCTCTTAATCTATGATCTGATTAATGAAACACA 3024
QY 2829 ACAACTGACAGATATTTGAAAGAAAGCTCTTTCATTTATGAGGCAAGTTACAGAG 2888
Db 3025 GCAGCTTACTCAAGATGCAAGTCCAGGCAATGGCTATCTCAACACTGTTACAGAG 3084
QY 2889 AGAATCTCTATCAGAGGAGATGCTCTTTCAGTGCCTTTGGG-----AATTATGA 2942
Db 3085 ACAGTTGACATCAAAACACTATGATGGCTCTACAGACCTTTGGGGAGCGATATGGAG 3144
QY 2943 CCTTCTGGAGCACTTGTGTGACGTTTGTTTAAGATGTTTCTTGAAGCCGATCC 3002
Db 3145 GAACGAGGCAACCTGGCTCACAGCTTGTCTGAAAGCTTTGGCCCAAGCTCGAGC 3204
QY 3003 TTACATGATATTTGATCAGATGATGTTTACAGAAACATACACTTGGCTTAAAGACATCA 3062
Db 3205 CTACATCTTCAATGATGAGACACATTTACCAAGCCCTCATATGAGCTCTCCAGAGCA 3264
QY 3063 GAAATCCACCGTGAATTTTGGATTCAGAGAGTGAATTCATGATGAGCTTCAAGGTGG 3122
Db 3265 GAAGGACAAATGGCTGTTTCAAGAGCTGGGTCACTGCTCAACATGCAATTAAGGGAGG 3324
QY 3123 CAATTAAGTCCAGTACACTTTACAGCTTATTTGTAATTTCTCTCTGGGATATGAAA 3182
Db 3325 AGTAGAAGATGAAGTGAACCTCTCTCCGCTATATACACATCCCTTCTGGAGATTCCTCT 3384
QY 3183 GTATCAGCTTAACATTTGATGCAAGAGTATCATTTTGGAGTGTGAATTCAGTAG 3242
Db 3385 CACAGTACTCACTGCTGTGTGCGCAATGCCCTGTTTGGCTGGAGTACGCTGGAAGAC 3444
QY 3243 AGGA 3246
Db 3445 AGCA 3448

RESULT 3
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```



```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-9300
; TELEFAX: (703) 683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: PTZpct-Fls
; US-08-232-463-14

```

```

Query Match      1.1%; Score 47.4; DB 1; Length 7218;
Best Local Similarity 4.5%; Pred. No. 0.008;
Matches 18; Conservative 217; Mismatches 168; Indels 0; Gaps 0;

```

```

QY 1149 AGAAGAAGAAATATGTAGTACATGACAGAGAAATCTACTGACTGAGG 1208
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1457 AAGAGATAGAAATTTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1398
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1209 CGGATTAACAGTGAATCAAGAAATGAGACTGTTCAAAATATTAATGCTGCC 1268
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1397 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1338
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1269 CCAAGTGAACCTTTAAGATTGAATCCCAATCTCGAGATTCAGTCACTACATT 1328
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1337 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1278
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1329 GAAGCCTATTCTCTGTAGTAAAGTAGCATGCGATTCCTTTAAGTCTCC 1388
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1277 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1218
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1389 TAGTAAGACATATCCAACTAAACAAAGATGAAATATTAAGTGGATCGCTTT 1448
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1217 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1158
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1449 TGAGTTGGTGTAGTGGCAACAAAGATTGAAGATTAAGCTATATGTAATCCAG 1508
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1157 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1098
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1509 GCGAAGATTGGTGGCTGTAGAAACAAATTCACCAATGTTCC 1551
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1097 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1055
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

RESULT 4
US-09-311-352B-1
; Sequence 1, Application US/09311352B
; Patent No. 6329500
; GENERAL INFORMATION:
; APPLICANT: Webb, Donna J.
; APPLICANT: Gonias, Steven L.
; TITLE OF INVENTION: Transforming Growth Factor-beta Binding Site
; FILE REFERENCE: 00370-02
; CURRENT APPLICATION NUMBER: US/09/311,352B
; CURRENT FILING DATE: 1999-05-13

```

```

; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 339
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-311-352B-1

```

```

Query Match      1.1%; Score 47.2; DB 4; Length 339;
Best Local Similarity 53.0%; Pred. No. 0.0011;
Matches 125; Conservative 0; Mismatches 108; Indels 3; Gaps 1;

```

```

QY 2056 GTCCGAAAGCATTTTCCAGAGACTGGATTGGCTAGACACCAACATGGGTTACAGATT 2115
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 97 GTACGAAAGTACTTCCCTGAGACATGATCTGGATTGGTGGTGAACCTCAGCAGG 156
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 2116 TACCAAGATTTGAAGTACTGTACTGATTTATCATCTTTGGGTGGCTACTGATTTT 2175
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 157 GTGGCTGAGAGTAGAGTAAACAGTCCCTGACACCATCACCGAGTGAAGGCAAGGCGCTTC 216
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 2176 GTGATCTGAGAGACTGGGTCTTGGACTAACACTACTCCAGTGGAGCTCCAGGCTTC 2235
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 217 TGCTGTCTGAGAGCTGAGACTGTGTATCTTTCACTGCTCT--CTCGAGGCTTC 273
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 2236 CAACCATTTTCTATTTTGTGATCTCTCTACTCTGTATCAGAGTGAAGATT 2291
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 274 CAGCCCTTCTTTGTGAGACTCAACATGCTTACTGTGATTCGTGAGAGGCTT 329
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

RESULT 5
US-08-447-411-1
; Sequence 1, Application US/08447411
; Patent No. 5773243

```

```

; GENERAL INFORMATION:
; APPLICANT: FRITZINGER, DAVID C.
; APPLICANT: BREDEHORST, REINHARD
; APPLICANT: VOGEL, CARL-WILHELM
; TITLE OF INVENTION: DNA ENCODING COBRA C3, CVP1, AND CVP2
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/447,411
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/043,747
; FILING DATE: 07-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5773243man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 1126-101-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5211 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown

```

MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 9..4961  
US-08-447-411-1

Query Match 1.1%; Score 46.4; DB 1; Length 5211;  
Best Local Similarity 50.4%; Pred. No. 0.012;  
Matches 113; Conservative 0; Mismatches 111; Indels 0; Gaps 0;

QY 3376 TCATCAGAGTCCAACTTTCTGACTCCTGGAGCAGCGCTCCGTGATTTGAAGTGCA 3435  
DB 3609 TCAACAGAGAGAAATCGTTGGAGAAATATATATCTGCACCCCTAATATTGAAGCACT 3668  
QY 3436 GCCTATGACTGCTCTGACACTTCTTCAATTTAGACTTGTAGAGGAAATCCCAATTATG 3495  
DB 3669 TCCATGCTCTTGTGGCCCTGCTGAAATGAAATTTGCTGAGGTGGCTCTGTATGTC 3728  
QY 3496 AGGTGGCTAAGCAGCAAGAAATAGCTTGGGTGGTTTGCATCTACTCAGATACACT 3555  
DB 3729 AGATGGCTGATAGATCAGAAATATTATGGGGGAAATATGGAACAACCCAGCAAGATT 3788  
QY 3556 GTGGCTTAAAGCTCTGTCTGATTTGACCCCTAATGATAC 3599  
DB 3789 ATGGTGTTCAGAGCTTGTGTAATATGAGATTGAGATGCTTAC 3832

RESULT 6  
US-09-241-606-5  
Sequence 5, Application US/09241606  
Patent No. 6472140  
GENERAL INFORMATION:  
APPLICANT: Tanzi, Rudolph E.  
APPLICANT: Kovacs, Dora  
APPLICANT: Saunders, Aleister J.  
TITLE OF INVENTION: Alpha-2-Macroglobulin Therapies and Drug Screening Methods for  
FILE REFERENCE: 0609.446003  
CURRENT APPLICATION NUMBER: US/09/241.606  
CURRENT FILING DATE: 1998-02-02  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 5  
LENGTH: 333  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(333)  
OTHER INFORMATION: A Binding Domain  
US-09-241-606-5

Query Match 1.0%; Score 44; DB 4; Length 333;  
Best Local Similarity 63.0%; Pred. No. 0.0092;  
Matches 68; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 3481 GGAATCCCAATTATGAGTGTCTAAGCAGCAAGAAATAGCTGGGTGTTGCATCT 3540  
DB 19 GCAACCAACATCGTAAGTGTGATCAGAGCAGCAAGAAATGCCCAGGGGGTTCTCTCC 78  
QY 3541 ACTCAGATACCACTGCTTTAAAGCTCTGTCTGATTTGACGCC 3588  
DB 79 ACCCAGACACAGTGTGCTCTCTCATGCTCTGTCCAAATATGAGCC 126

RESULT 7  
US-09-241-606-3  
Sequence 3, Application US/09241606  
Patent No. 6472140  
GENERAL INFORMATION:  
APPLICANT: Tanzi, Rudolph E.  
APPLICANT: Kovacs, Dora  
APPLICANT: Saunders, Aleister J.

TITLE OF INVENTION: Alpha-2-Macroglobulin Therapies and Drug Screening Methods for  
FILE REFERENCE: 0609.446003  
CURRENT APPLICATION NUMBER: US/09/241.606  
CURRENT FILING DATE: 1998-02-02  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 3  
LENGTH: 750  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(750)  
OTHER INFORMATION: A /LRP Binding Domain  
US-09-241-606-3

Query Match 1.0%; Score 44; DB 4; Length 750;  
Best Local Similarity 63.0%; Pred. No. 0.016;  
Matches 68; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 3481 GGAATCCCAATTATGAGTGTCTAAGCAGCAAGAAATAGCTGGGTGTTGCATCT 3540  
DB 19 GCAACCAACATCGTAAGTGTGATCAGAGCAGCAAGATGCCAGGGGGTTCTCTCC 78  
QY 3541 ACTCAGATACCACTGCTTTAAAGCTCTGTCTGATTTGACGCC 3588  
DB 79 ACCCAGACACAGTGTGCTCTCATGCTCTGTCCAAATATGAGCC 126

RESULT 8  
US-09-356-952-11  
Sequence 11, Application US/09356952  
Patent No. 6117663  
GENERAL INFORMATION:  
APPLICANT: Borlack-Sjodin, Ann  
APPLICANT: Margalit, S. M.  
APPLICANT: Bor-Sogli, Dafna  
APPLICANT: Cole, Philip  
APPLICANT: Kurliyan, John  
TITLE OF INVENTION: A CRYSTAL OF A RAS-SOS COMPLEX AND METHODS OF USE  
FILE REFERENCE: 600-1-228N  
CURRENT APPLICATION NUMBER: US/09/356.952  
CURRENT FILING DATE: 1999-07-19  
EARLIER APPLICATION NUMBER: 60/093.631  
EARLIER FILING DATE: 1998-07-21  
NUMBER OF SEQ ID NOS: 14  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 11  
LENGTH: 5398  
TYPE: DNA  
ORGANISM: Saccharomyces cerevisiae  
US-09-356-952-11

Query Match 1.0%; Score 41.4; DB 3; Length 5398;  
Best Local Similarity 47.2%; Pred. No. 0.37;  
Matches 126; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

QY 802 TTTTACCTTTATCTTTTGGGAGAAAGAAATAATTAACAAAATTAAAGATAAT 861  
DB 2671 TTTTACCTTTTCAACGCTCTCTCATTAACAGAACTTTTCAAGACAGATTCAAAA 2730  
QY 862 GGAATGCAAACTTCTTTTATGATGAAGATGAAAATGTAATGATTTCTCAAT 921  
DB 2731 AGGAGAGAAAAATATCATTAATCTAGACACTTTGAAATCAATGAAGAAATCTCG 2790  
QY 922 GGAATTTGATACCGATGATCTATCTCCCTGACCACTAGAAATTTTAAACAGAGTG 981  
DB 2791 CAATTTTAAATAATTAATATGCTACAGGTGAACCTTTAAATTAATTAAGTAAACC 2850  
QY 982 ACAGATCAAGTACAGTATTTCAAGAAATGTAAGCACTAATGTGTTCTTCAAGCAAT 1041

```

Db      2851 AAAAGCAGATTGAGAAATTGGAATTAATTCAGACATACGACAAATTAATCAGAAAT 2910
Oy      1042 GATTACATCATTTGAGTTTGTGATTAT 1068
        ||| ||| ||| ||| |||
Db      2911 GTTTACTATTGTGAGTACTGAGAAAT 2937

RESULT 9
US-08-916-421B-1/c
; Sequence 1, Application US/08916421B
; Patent No. 6503729
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanoco
; Patent No. 6503729
; FILE REFERENCE: PB275
; CURRENT APPLICATION NUMBER: US/08/916,421B
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/024,428
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98150)..(98150)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98266)..(98266)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98343)..(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (10398)..(10398)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (148548)..(148548)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (16385)..(16385)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (19189)..(19189)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (19195)..(19195)
; OTHER INFORMATION: n equals a, t, c, or g

```

```

; NAME/KEY: misc_feature
; LOCATION: (231980)..(231980)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (234187)..(234187)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (234220)..(234220)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (234814)..(234814)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (309398)..(309398)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (309418)..(309418)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (312837)..(312837)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (312993)..(312993)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (319226)..(319226)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (559167)..(559167)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (559241)..(559241)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (600992)..(600992)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (622708)..(622708)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (657081)..(657081)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (657203)..(657203)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (674435)..(674435)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (682442)..(682442)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (713652)..(713652)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (741684)..(741684)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (779455)..(779455)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (779676)..(779676)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (855539)..(855539)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (871619)..(871619)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1084830)..(1084830)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature

```

LOCATION: (1096846)..(1096846)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1115881)..(1119881)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1130881)..(1130881)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1310988)..(1310988)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1313224)..(1313224)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1349491)..(1349491)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1470091)..(1470091)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1569020)..(1569020)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1602912)..(1602912)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1603734)..(1603734)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1637998)..(1637998)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1664854)..(1664854)  
OTHER INFORMATION: n equals a, t, c, or g  
US-08-916-421B-1

Query Match 1.0%; Score 41.4; DB 4; Length 1664976;  
Best Local Similarity 52.0%; Pred. No. 19;  
Matches 93; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

Qy 3725 TTGTATTAATGTAAGGCTTCTGGCTCTTGAAGACGAAGATCTATCCAAATCAG 3784  
Db 539968 TTGAGCTACTGATGATGCTTACATCTATGCAACGGAACAATGATTAAGTGCCAA 539909  
Qy 3785 AAGCTTTGATTGATGCTGCTGTAAGAAATTAAGATGATCTCAATCATGTCGATT 3844  
Db 539908 AAACATTAAGGTATGATTAATGTAAGAAATTAAGATGATCTCAATCATGTCGATT 539849  
Qy 3845 TGAATGTGTGTAACAAGCTTTTCGGCCCGGCTAGAGTGCATGCTCTTAAGAAATT 3903  
Db 539848 TAAGGCTTTGTAAGAAATGGAAGAGAGAGACAATCATATGCTATGATGATGCT 539790

RESULT 10  
US-09-601-198-80  
Sequence 80, Application US/09601198  
Patent No. 6531583  
GENERAL INFORMATION:  
APPLICANT: Casseil, Gail H.  
APPLICANT: Chen, Elison Y.  
APPLICANT: Glass, Jennifer S.  
APPLICANT: Glass, John I.  
APPLICANT: Heiner, Cheryl R.  
APPLICANT: Lefkowitz, Elliot  
TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREAPLASMA  
FILE REFERENCE: UAB-13452/22  
CURRENT APPLICATION NUMBER: US/09/601,198  
CURRENT FILING DATE: 2000-12-08

PRIOR APPLICATION NUMBER: 60/073,189  
PRIOR FILING DATE: 1998-01-30  
NUMBER OF SEQ ID NOS: 181  
SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO 80  
LENGTH: 1581  
TYPE: DNA  
ORGANISM: Ureaplasma urealyticum  
US-09-601-198-80

Query Match 0.9%; Score 40; DB 4; Length 1581;  
Best Local Similarity 43.8%; Pred. No. 0.4;  
Matches 175; Conservative 0; Mismatches 225; Indels 0; Gaps 0;

Qy 3542 CTGAGATACACAGTGTGCTTAAAGGCTGCTGTAATTCAGCCCTATGATGATCAG 3601  
Db 578 CTATGATGACATGATCTTTTGAATCTGAAATGCTTTGATGATCTATTTGATCTCA 637  
Qy 3602 AAAGACAAATATCCAAAGTACCGGCGCTTACCTCAACAGTCTTCTGCTG 3661  
Db 638 AAGTAATCTGTTGAAAGGCTTTAGCATCTAAATTAACAAATTCATTTCAAAA 697  
Qy 3662 TACAGCCATGACGATTAATTTCCGCAATGCTTTGATTTGCTATTTGACGCTCA 3721  
Db 698 CAGCAGTAAAGTGTGTTACATTTATGATGATGCTATGCTTAAATTTATGATAGA 757  
Qy 3722 ATGTTGATTAATGTAAGGCTTCTGGCTCTTGAAGACGAAGATCTATCCAAATC 3781  
Db 758 ATGCAATCTTAATTTAAATTAACAGAAATTTGTAATGATTAATTAACCAAAAT 817  
Qy 3782 AAGAGCTTTGATTTGATGCTGCTGTAAGAAATTAAGATGATCTCAATCATGTCG 3841  
Db 818 ATAAAGTTATTTAGCTGCTGGAATTTTAAATTAATTAATTAATTAATTTATCTT 877  
Qy 3842 ATTTGATGCTGTAACAAGCTTTTCGGCCCGGCTAGAGTGCATGCTCTTATGCAAG 3901  
Db 878 TACCAATGATGATTTGTAACGCAAGTATGATGATGATGATGATGATGATGAT 937  
Qy 3902 TTAACCTATTAAGGCTTATGCTTATGCTTCAAGAAAT 3941  
Db 938 ATACTTATCAAAAGCTTTTGAATGCAAAAGATGAT 977

RESULT 11  
US-09-601-198-60/c  
Sequence 60, Application US/09601198  
Patent No. 6531583  
GENERAL INFORMATION:  
APPLICANT: Casseil, Gail H.  
APPLICANT: Chen, Elison Y.  
APPLICANT: Glass, Jennifer S.  
APPLICANT: Glass, John I.  
APPLICANT: Heiner, Cheryl R.  
APPLICANT: Lefkowitz, Elliot  
TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREAPLASMA  
FILE REFERENCE: UAB-13452/22  
CURRENT APPLICATION NUMBER: US/09/601,198  
CURRENT FILING DATE: 2000-12-08  
PRIOR APPLICATION NUMBER: 60/073,189  
PRIOR FILING DATE: 1998-01-30  
NUMBER OF SEQ ID NOS: 181  
SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO 60  
LENGTH: 15016  
TYPE: DNA  
ORGANISM: Ureaplasma urealyticum  
US-09-601-198-60

Query Match 0.9%; Score 39.4; DB 4; Length 15016;  
Best Local Similarity 48.1%; Pred. No. 2.9;  
Matches 112; Conservative 0; Mismatches 121; Indels 0; Gaps 0;

QY 703 ACACCTATATGTTCTATGATTTAAGCTTAATGTAACATCAAGCAAGTAT 762  
DB 2859 ACACCTATATGATGATATTAATAAAGATATTCAGCAAAATATTCAGTATTT 2800  
QY 763 ACATATGGGAAGCCAGTGAAGAGCGTAACGCTTACCTTTTCTTACCTTTG 822  
DB 2799 GAACTGATACACATTAATGAAATTTTAATGAGCAATGTAATCTTTTAAAT 2740  
QY 823 GGAAGAAAGAAATATTTACAAATTCATTAAGATAATGATGCAAACTTCTTTT 882  
DB 2739 AATTAGAGCCTAATATTTAATAATAAATTAATAATGATGAGCAAAACCACT 2680  
QY 883 AATGATGAAGATGAATAAATGTAATGATTTCTTAATGATGACTTTGATA 935  
DB 2679 AAGGCTATGCTAATTTAATTCAAAGCTAATAACATGATTAATGATAA 2627

RESULT 12  
US-09-350-756-9  
; Sequence 9, Application US/09350756  
; Patent No. 6495143  
; GENERAL INFORMATION:  
; APPLICANT: U.S. Army Medical Research Institute for Infectious Diseases  
; APPLICANT: John S. Lee  
; APPLICANT: Peter Pushko  
; APPLICANT: Michael D. Parker  
; APPLICANT: Jonathan F. Smith  
; APPLICANT: Mark T. Dertzbaugh  
; APPLICANT: Leonard Smith  
; TITLE OF INVENTION: Botulinum Neurotoxin Vaccine  
; FILE REFERENCE: 003/124/SAP R1ID 98-21  
; CURRENT APPLICATION NUMBER: US/09/350,756  
; EARLIER FILING DATE: 1999-07-09  
; EARLIER APPLICATION NUMBER: US 60/092,416  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: Apple Macintosh Microsoft Word 6.0  
; SEQ ID NO 9  
; LENGTH: 1327  
; TYPE: DNA  
; ORGANISM: Clostridium botulinum  
; FEATURE:  
US-09-350-756-9

Query Match 0.9%; Score 38.8; DB 4; Length 1327;  
Best Local Similarity 49.5%; Pred. No. 0.79;  
Matches 100; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

QY 1472 AACGATTTGAAGGATTAAGCTATATGTAATCCAGGGAGAGTTGGCTGTAGGA 1531  
DB 801 AATATATGATGATATTTAGAGTATATGATCTTAAAGGCGCTAGAGGCTAATGAC 860  
QY 1532 AACAAATTCACAAATGTTCTCTTTAACAACGAAATTTCTGACTCCAAAGCCTGTG 1591  
DB 861 TACAATGACTCAAAACATTTATTTAAATTCAGTTTGTATAGGGGCAAAATTTATTA 920  
QY 1592 TTAATGATATATATGTAAGTATGATGGGAATTAATGATGTTCTAAATAATTCG 1651  
DB 921 TAAAAAATATCTCTCTGGAATAAAGATTAATTTGTTGAATAATGATGCTGATATA 980  
QY 1652 TTCAGCTGTTTAAAAATAA 1673  
DB 981 TTAATGATGATTAATAATAA 1002

RESULT 13  
US-08-955-138-2/c  
; Sequence 2, Application US/08955138A  
; Patent No. 5977435  
; GENERAL INFORMATION:  
; APPLICANT: Lefebvre, Daniel D.  
; APPLICANT: Gelliaty, Kevin S.  
; TITLE OF INVENTION: PLANT PHOSPHATASES

; FILE REFERENCE: PPL97-01  
; CURRENT APPLICATION NUMBER: US/08/955,138A  
; CURRENT FILING DATE: 1997-10-21  
; NUMBER OF SEQ ID NOS: 119  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 3981  
; TYPE: DNA  
; ORGANISM: SOLANUM TUBEROSUM  
US-08-955-138-2

Query Match 0.9%; Score 38.8; DB 2; Length 3981;  
Best Local Similarity 50.0%; Pred. No. 1.7;  
Matches 97; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

QY 1503 ATCCAGGGAGACAGTTGGCTGTAGGAACAATTCACATGTTCTTTAACC 1562  
DB 3066 ATCTAGACACAAATTTGTTAAATTTGACACCTTAATTAATTTCTCAGAGTTAACTT 3007  
QY 1563 AGAAATTTCTGGACTCCAAAGCCTGTGTAATTTGTAATTAATGAAGATGGGGA 1622  
DB 3006 AGAATACCTTAACCGCCTACTAAGTATATTTGTTATATGAAGCTAGCAAGTA 2947  
QY 1623 AATTATAGTATGTTCTAAATTTCTGTTCACTTTGTTTAAAAATTAAGATPAACT 1682  
DB 2946 AATTATTTTGTGTTTAACTTTTGTGAATCATTAATTAATTAAGTAAATGTT 2887  
QY 1683 AATTGGAGTAAAG 1696  
DB 2886 CAATTTGTTGAAG 2873

RESULT 14  
US-09-071-035-83  
; Sequence 83, Application US/09071035  
; Patent No. 6448043  
; GENERAL INFORMATION:  
; APPLICANT: Gil H. Choi  
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides  
; NUMBER OF SEQUENCES: 496  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/071,035  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: A. Anders Brooks  
; REGISTRATION NUMBER: 36,373  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 83:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1579 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
US-09-071-035-83

Query Match	0.9%	Score 38.6	DB 4	Length 1579
Best Local Similarity	49.3%	Pred. No. 1		
Matches 101	Conservative 0	Mismatches 104	Indels 0	Gaps 0
Qy	1501	GTATCCAGGGGACAGTTGGTGGCTGTAAGAAAAAATAATTCAGAAATGTTCTCTTTAACA	1560	
Db	188	GTACCAAAAGCATTTGCCGATGATTTCCGAAGATGAGAAAACTACACGATTTCTTTGACA	247	
Qy	1561	CCAGAAATTTCTTGACCTCCAAAGACCTGATATGTGTAATTAATTAATTAAGATGATGGG	1620	
Db	248	AAAGACGCGTTTGAGATTAACGATGATCTGTCCACGACATGATTTGAAATATGCTTGG	307	
Qy	1621	GAAATTAATGAATGATGTTCTTAAAAATTCCTGTTACGCTGTTTTTAAAAATAGATTAAG	1680	
Db	308	AAAAAATGATGATGCTTAAAAAACGCGCTTTGTTATAGTCTCTCATCGTGAACAATTT	367	
Qy	1681	CTATATGAGATTAAGTGAAGCTG	1705	
Db	368	CAAATATGTCAGAAATCTCAGCGG	392	

```

RESULT 15
US-09-071-035-81
: Sequence 81, Application US/09071035
: Patent No. 6448043
: GENERAL INFORMATION:
: APPLICANT: Gil H. Choi
: TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
: NUMBER OF SEQUENCES: 496
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Human Genome Sciences, Inc.
: STREET: 9410 Key West Avenue
: CITY: Rockville
: STATE: Maryland
: COUNTRY: USA
: ZIP: 20850
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
: COMPUTER: HP Vectra 486/33
: OPERATING SYSTEM: MSDOS version 6.2
: SOFTWARE: ASCII Text
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/071,035
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: A. Anders Brookes
: REGISTRATION NUMBER: 36,373
: REFERENCE/DOCKET NUMBER: PB369P2
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (301) 309-8504
: TELEFAX: (301) 309-8512
: INFORMATION FOR SEQ ID NO: 81:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1680 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
:
US-09-071-035-81

Query Match 0.9%; Score 38.6; DB 4; Length 1680;
Best Local Similarity 49.3%; Pred. No. 1.1; 104; Indels 0; Gaps 0;
Matches 101; Conservative 0; Mismatches 0;

QY 1501 GTATCCAGGGAGACGTTGTGTGGCTGTATGAGAAAACAAATTCACAATGTTCTCTTTAACA 1560
D 286 GTTGCAGAAACGATTCGCCGATGATTTTCAGAGATGGAACCACTACACGATTTCTTTGAGA 345
Y 1561 CCAAGAAATTTCTTGACTCCAAAGCCTGTGTATATGTTGTAATTAATGAATGATGCG 1620

```

Db 346 A A A G A G C G G T T T G A G T A A C G A T C C T C A C A C A C A T G A T T T T G A A T A T G C T T G 405

Qy 1621 G A A A T T A T A G T A G T G T T C T A A A A A T T C T G T C A G C T T G T T T T A A A A T A G A T A A G 1680

Db 406 A A A A A A A T G A T T G A T T C C T A A A A A A G G G C T T G T T A T A G C T T C C A T C G T T G A A A C A A T T 465

Qy 1681 C T A T A T T G G A G T A A A G T G A A A C T G 1705

Db 466 C A A A A T G T G C A G A A A T C T C A G C G G 490

Search completed: January 16, 2004, 21:01:59  
Job time : 274 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2004, 14:38:51 ; Search time 1338 Seconds

(without alignments)  
11293.468 Million cell updates/sec

Title: US-10-020-095-3

Perfect score: 4287  
Sequence: 1 atgcaggagccaccgcctcctc.....ttatggaacttgctgtga 4287

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 2324096 seqs, 1762381658 residues

Total number of hits satisfying chosen parameters: 4648192

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09C\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 17: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 18: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2584.6	60.3	3033	10	US-09-833-381-1810 Sequence 1810, App
2	1969.4	45.9	2273	12	US-10-108-260A-953 Sequence 953, App
3	730.8	17.0	1300	13	US-10-133-013-223 Sequence 223, App
4	541.6	12.6	875	13	US-10-160-162-48 Sequence 48, Appl
5	541.6	12.6	875	13	US-09-820-649-48 Sequence 48, Appl
6	272	6.3	354	10	US-09-960-352-12867 Sequence 12867, A
7	194.4	4.5	689	13	US-10-027-632-204326 Sequence 204326, Sequence 266, App
8	194.4	4.5	689	14	US-10-027-632-204326 Sequence 204326, Sequence 266, App
9	130	3.0	4677	13	US-10-316-253-266 Sequence 266, App
10	130	3.0	4701	13	US-10-316-253-268 Sequence 268, App
11	126.8	3.0	4615	10	US-09-880-107-3727 Sequence 3727, App
12	126.8	3.0	4615	10	US-10-006-285-405 Sequence 405, App
13	125.2	2.9	4488	13	US-09-981-151A-9 Sequence 9, Appl
14	119.8	2.8	4501	12	US-10-094-886-125 Sequence 125, App
15	119.4	2.8	1140	9	US-09-764-853-362 Sequence 362, App

16	119.4	2.8	1140	11	US-09-764-876-17 Sequence 17, Appl
17	118.6	2.8	2092	11	US-09-764-876-12 Sequence 12, Appl
18	117.8	2.7	4595	10	US-09-917-800A-1531 Sequence 1531, App
19	117.4	2.7	2092	9	US-09-764-853-163 Sequence 163, App
20	116.4	2.7	2252	12	US-10-108-260A-1292 Sequence 1292, App
21	114.2	2.7	4527	13	US-09-756-247-3 Sequence 3, Appl
22	114.2	2.7	4527	13	US-09-756-247-5 Sequence 5, Appl
23	110	2.6	2041	13	US-10-292-081A-4 Sequence 4, Appl
24	110	2.6	2041	13	US-10-087-188-1 Sequence 1, Appl
25	110	2.6	4021	9	US-09-925-301-552 Sequence 552, App
26	110	2.6	4422	9	US-09-873-403-4 Sequence 4, Appl
27	110	2.6	4530	13	US-10-292-081A-2 Sequence 2, Appl
28	110	2.6	4576	13	US-10-292-081A-7 Sequence 7, Appl
29	110	2.6	4577	9	US-09-873-403-3 Sequence 3, Appl
30	110	2.6	4577	10	US-09-880-107-2236 Sequence 2236, App
31	110	2.6	4577	12	US-10-331-996A-19 Sequence 19, Appl
32	110	2.6	4577	13	US-09-873-319-408 Sequence 408, App
33	110	2.6	4577	13	US-09-960-706-654 Sequence 654, App
34	110	2.6	4577	13	US-10-292-081A-3 Sequence 3, Appl
35	110	2.6	4577	13	US-10-292-081A-5 Sequence 5, Appl
36	110	2.6	4577	14	US-10-052-817-1 Sequence 1, Appl
37	110	2.6	4577	15	US-10-076-816-56 Sequence 56, Appl
38	110	2.6	4809	13	US-10-240-865-178 Sequence 178, App
39	110	2.6	4823	13	US-09-971-392-172 Sequence 172, App
40	110	2.6	5092	13	US-10-006-285-404 Sequence 404, App
41	110	2.6	5092	13	US-09-971-429B-2 Sequence 2, Appl
42	106.4	2.5	4771	13	US-09-756-247-27 Sequence 27, Appl
43	99	2.3	4426	12	US-10-094-886-123 Sequence 123, App
44	93.4	2.2	451	11	US-09-918-995-12615 Sequence 12615, A
45	85.6	2.0	5191	13	US-10-006-285-206 Sequence 206, App

#### ALIGNMENTS

RESULT 1  
US-09-833-381-1810  
; Sequence 1810, Application US/09833381  
; Patent No. US20020132090A1  
; GENERAL INFORMATION:  
; APPLICANT: Robison, Keith E.  
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs  
; FILE REFERENCE: 5800-119  
; CURRENT APPLICATION NUMBER: US/09/833.381  
; PRIOR FILING DATE: 2001-04-11  
; PRIOR APPLICATION NUMBER: 09/516,448  
; NUMBER OF SEQ ID NOS: 2050  
; SOFTWARE: FASTSEQ for Windows Version 3.0  
; SEQ ID NO 1810  
; LENGTH: 3033  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-833-381-1810

Query Match	Best Local Similarity	Score	2584.6;	DB 10;	Length 3033;
Matches 2648;	Conservative 0;	Mismatches 4;	Indels 51;	Gaps 1;	
QY	1636	GTCTAAATTCCTGTTACCTGTTTAAATAGATAAGTATATGGAGTAA	1635		
DB	6	GTCTGAAATTCCTGTTACCTGTTTAAATAGATAAGTATATGGAGTAA	65		
QY	1656	GTGAAAGTGAACCATCTGAGAAAGTCTCTTNGATCTGTGACACAGCTGACTCC	1755		
DB	66	GTGAAAGTGAACCATCTGAGAAAGTCTCTTNGATCTGTGACACAGCTGACTCC	125		
QY	1756	ATAGTGGATTTAGCGTTGACAAAGTGTGAATGTGATGCTTAAGATTT	1815		
DB	126	ATAGTGGATTTAGCGTTGACAAAGTGTGAATGTGATGCTTAAGATTT	185		
QY	1816	ACAATGAAATGTGTCATGATGGAATTTATACACAGATATTTATGGCATG	1875		

Db 186 ACAATGAAAATGTCCTCATGATGTTGCAACTTATTAACAAGATATTTATTAAGGCATG 245  
 Qy 1876 TTCAATGAATCTTTTTCAGAGTCTTTTCAGAGATGAGACTCTGGATTAATGACAGATGCAAAAC 1935  
 Db 246 TTCAATGAATCTTTTTCAGAGTCTTTTCAGAGATGAGACTCTGGATTAATGACAGATGCAAAAC 305  
 Qy 1936 CTCACGAAGATTAATTTGATGCTTTTATGCAATGCAAAATATGCTGAAGGTTTATG 1995  
 Db 306 CTCACGAAGATTAATTTGATGCTTTTATGCAATGCAAAATATGCTGAAGGTTTATG 365  
 Qy 1996 GAGGAATATGAAGGATATTTGATGATTAATGCACTTTCTTTGGGATGAGAGTCCAT 2055  
 Db 366 GAGGAATATGAAGGATATTTGATGATTAATGCACTTTCTTTGGGATGAGAGTCCAT 425  
 Qy 2056 GTCCGAAGACATTTTTCAGAGACTTTGATTTGGCTGACACCAACATGGGTTTACAGAT 2115  
 Db 426 GTCCGAAGACATTTTTCAGAGACTTTGATTTGGCTGACACCAACATGGGTTTACAGAT 485  
 Qy 2116 TACCAAGATTTGAAGTATCTGATCTGATTTCTATCACTTTGGGCTGATCTGCTTTT 2175  
 Db 486 TACCAAGATTTGAAGTATCTGATCTGATTTCTATCACTTTGGGCTGATCTGCTTTT 545  
 Qy 2176 GTGATCTGAGGACCTGGGCTTTGGAATAAATGCACTTCCAGTGAAGCTCCAGGCTTC 2235  
 Db 546 GTGATCTGAGGACCTGGGCTTTGGAATAAATGCACTTCCAGTGAAGCTCCAGGCTTC 605  
 Qy 2236 CAACCATTTTTCATTTTTCATTTTTCATTTTTCATTTTTCATTTTTCATTTTTCATTTT 2295  
 Db 606 CAACCATTTTTCATTTTTCATTTTTCATTTTTCATTTTTCATTTTTCATTTTTCATTTT 665  
 Qy 2296 TTGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2355  
 Db 666 TTGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 725  
 Qy 2356 AAAAGTGAATTTGATTAATTTGATTAATTTGATTAATTTGATTAATTTGATTAATTTGAT 2415  
 Db 726 AAAAGTGAATTTGATTAATTTGATTAATTTGATTAATTTGATTAATTTGATTAATTTGAT 785  
 Qy 2416 CAGACCTCTGCTGCTCCAGTGAAGATGAGGCACTGCTTTTCCATCAGGCAACA 2475  
 Db 786 CAGACCTCTGCTGCTCCAGTGAAGATGAGGCACTGCTTTTCCATCAGGCAACA 845  
 Qy 2476 CATCTGGGAAATTTCTATCACTGACAGCTCTTTTCAACCACTGCTTCTGATGCTGTC 2535  
 Db 846 CATCTGGGAAATTTCTATCACTGACAGCTCTTTTCAACCACTGCTTCTGATGCTATC 905  
 Qy 2536 ACCGAGATTTTATGTAAGGCTGAAGGAATGAAGAAATCATATTCATTCATCTTA 2595  
 Db 906 ACCGAGATTTTATGTAAGGCTGAAGGAATGAAGAAATCATATTCATTCATCTTA 965  
 Qy 2596 TTAGACTTGAATGATGAGCTACAGAGTACCTGAAACCTTTGAGTTTCTCATTTTCT 2655  
 Db 966 TTAGACTTGAATGATGAGCTACAGAGTACCTGAAACCTTTGAGTTTCTCATTTTCT 1025  
 Qy 2656 CCTATATCAGTGAAGTGAAGGCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2715  
 Db 1026 CCTATATCAGTGAAGTGAAGGCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1085  
 Qy 2716 CCTTTCATCAATGCTTACCTCATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 2775  
 Db 1086 CCTTTCATCAATGCTTACCTCATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 1145  
 Qy 2776 ATGATTAATTTTGTCCAAATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 2835  
 Db 1146 ATGATTAATTTTGTCCAAATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1205  
 Qy 2836 ACAGATTAATTTGAAAGAAAGTCTTTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 2895  
 Db 1206 ACAGATTAATTTGAAAGAAAGTCTTTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1265  
 Qy 2896 CTCTATCAGAGGAAAGATGCTTTTCACTGCTTTTGGGAAATTAATGACCTTTCTGGAGC 2955  
 Db 1266 CTCTATCAGAGGAAAGATGCTTTTCACTGCTTTTGGGAAATTAATGACCTTTCTGGAGC 1325

Qy 2956 ACTTGTTGTCAGCTTTTGTTTTAAAGTGTTCCTTGAAAGCCGATCTTATCATAGATAT 3015  
 Db 1326 ACTTGTTGTCAGCTTTTGTTTTAAAGTGTTCCTTGAAAGCCGATCTTATCATAGATAT 1385  
 Qy 3016 GATCAGAATGTGTTACACAGATATCATCTTGGCTTAAAGACATCAGAAATCCAGCT 3075  
 Db 1386 GATCAGAATGTGTTACACAGATATCATCTTGGCTTAAAGACATCAGAAATCCAGCT 1445  
 Qy 3076 GAATTTGGGATCAGAGAGATGATTCATGATGAGCTTCAGGTGGCAATTAAGTCCA 3135  
 Db 1446 GAATTTGGGATCAGAGAGATGATTCATGATGAGCTTCAGGTGGCAATTAAGTCCA 1505  
 Qy 3136 GTAAACCTTACAGCTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 3195  
 Db 1506 GTAAACCTTACAGCTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1565  
 Qy 3196 ATGATGTGCAAGATATTCATTTTGGAGTCTGAAATTCAGATGAGAAATTTCAAG 3255  
 Db 1566 ATGATGTGCAAGATATTCATTTTGGAGTCTGAAATTCAGATGAGAAATTTCAAG 1625  
 Qy 3256 AATTATCTAGGCTTATTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 3315  
 Db 1626 AATTATCTAGGCTTATTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 3375  
 Qy 3316 GAAGCTTTGAATATGCTGATCTTGAGAGCAAGAAAGAGAGTGCATTTCTGGGTG 3375  
 Db 1686 GAAGCTTTGAATATGCTGATCTTGAGAGCAAGAAAGAGAGTGCATTTCTGGGTG 1745  
 Qy 3376 TCATCAGATTCAAATTTTCTGACTCTGGAGCCAGCTCTCTGGATTTGAAGTTGCA 3435  
 Db 1746 TCATCAGATTCAAATTTTCTGACTCTGGAGCCAGCTCTCTGGATTTGAAGTTGCA 1805  
 Qy 3436 GCCATGCACTGCTCAGCTTCTTCAATTTCAAGCTTCAAGGGAATCCCAATTATG 3495  
 Db 1806 GCCATGCACTGCTCAGCTTCTTCAATTTCAAGCTTCAAGGGAATCCCAATTATG 1865  
 Qy 3496 AGGTGCTAAGCAGGCAAGAAATAGCTTGGTGTGTTTTCATCTCAGATACCACT 3555  
 Db 1866 AGGTGCTAAGCAGGCAAGAAATAGCTTGGTGTGTTTTCATCTCAGATACCACT 1925  
 Qy 3556 GTGCTTTAAAGGCTCTGCTGATTTGACAGCTTATGAATACAGAAAGACAAATATC 3615  
 Db 1926 GTGCTTTAAAGGCTCTGCTGATTTGACAGCTTATGAATACAGAAAGACAAATATC 1985  
 Qy 3616 CAAGTACGCTGAGGGGCTTACCTCAAGTCTT----- 3651  
 Db 1986 CAAGTACGCTGAGGGGCTTACCTCAAGTCTT----- 2045  
 Qy 3652 -----CTTGTGTGTGTAAGGCAATGCGATTAATAT 3684  
 Db 2046 AACGGCTTACTCTTACAGACAGAGCTTGTGTGTGTAAGGCAATGCGATTAATAT 2105  
 Qy 3685 TCCGCAATGTTTGGATTTGCTATTTGTCACTCAATGTTGATTAATGGAAGCT 3744  
 Db 2106 TCCGCAATGTTTGGATTTGCTATTTGTCACTCAATGTTGATTAATGGAAGCT 2165  
 Qy 3745 TCTGGCTTCTTGAAGACGAAGATCATTCGCAAAATCAAGAAAGCTTTGATTAAGTGT 3804  
 Db 2166 TCTGGCTTCTTGAAGACGAAGATCATTCGCAAAATCAAGAAAGCTTTGATTAAGTGT 2225  
 Qy 3805 GCTGTAAAGAAATTAAGATGATCTCAATCATGTGATTTGAATGTGTACAGCTTT 3864  
 Db 2226 GCTGTAAAGAAATTAAGATGATCTCAATCATGTGATTTGAATGTGTGTACAGCTTT 2285  
 Qy 3865 TCGGGCCCGGGTGAAGTGCATGGCTTTATGAAAGTTAACTATTAATGAGCTTTATG 3924  
 Db 2286 TCGGGCCCGGGTGAAGTGCATGGCTTTATGAAAGTTAACTATTAATGAGCTTTATG 2345  
 Qy 3925 GTGCTTTCAGAGCAATTTCTGAGCGAGACAGTGAAGAAATGGAATTTGATCATGCA 3984  
 Db 2346 GTGCTTTCAGAGCAATTTCTGAGCGAGACAGTGAAGAAATGGAATTTGATCATGCA 2405



QY 3985 AAACCTCACTCTATTAGATTCTGTAAATGAACCCAGTTTGTGTAAATATCTCTGCT 4044  
 Db 2406 AAACCTCACTCTATTAGATTCTGTAAATGAACCCAGTTTGTGTAAATATCTCTGCT 2465  
 QY 4045 GTGAGAACTTTAAGTTTCAATACCAAGATGCTTCATAGTGAATTAATAT 4104  
 Db 2466 GTGAGAACTTTAAGTTTCAATACCAAGATGCTTCATAGTGAATTAATAT 2525  
 QY 4105 GAGCCAGAGAGACAGCGGTGAGAGATTACACTCTGAAGTGAAGCTGTCTCTGAGAC 4164  
 Db 2526 GAGCCAGAGAGAGACAGCGGTGAGAGATTACACTCTGAAGTGAAGCTGTCTCTGAGAC 2585  
 QY 4165 CTTTGAGTGAATGCCAGGCTGCTCTTGTGAGATGAGCTTCAAGCTCCATCAT 4224  
 Db 2586 CTTTGAGTGAATGCCAGGCTGCTCTTGTGAGATGAGCTTCAAGCTCCATCAT 2645  
 QY 4225 CACTCTTCACTATTTTATTTTCTGTTCAGACTTCTGACTTATGAACTTTGGCTG 4284  
 Db 2646 CACTCTTCACTATTTTATTTTCTGTTCAGACTTCTGACTTATGAACTTTGGCTG 2705  
 QY 4285 TGA 4287  
 Db 2706 TGA 2708

RESULT 2  
 US-10-108-260A-953  
 ; Sequence 953, Application US/10108260A  
 ; Publication No. US20040005560A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: HELIX RESEARCH INSTITUTE  
 ; TITLE OF INVENTION: No. US20040005560A1e1 full length cDNA  
 ; FILE REFERENCE: H1-A0106  
 ; CURRENT APPLICATION NUMBER: US/10/108, 260A  
 ; CURRENT FILING DATE: 2002-03-27  
 ; NUMBER OF SEQ ID NOS: 5458  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 953  
 ; LENGTH: 2273  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; US-10-108-260A-953

Query Match 45.9%; Score 1969.4; DB 12; Length 2273;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1970; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAGAGGCGCCACCGCTCTGACCGCGCCACCTCTCTGCGTGTGACCGCGCGCTG 60  
 Db .66 ATGAGAGGCGCCACCGCTCTGACCGCGCCACCTCTCTGCGTGTGACCGCGCGCTG 125  
 QY 61 GCCGTGGCTCCCGGGCTCGGTTCCTGTGACAGCCCGAGATCATGAGCCCGAGGA 120  
 Db 126 GCCGTGGCTCCCGGGCTCGGTTCCTGTGACAGCCCGAGATCATGAGCCCGAGGA 185  
 QY 121 AATGTACTATTGGGGTGAAGCTTCTGAAACACTGACCTTCAAGATGACTGGAAGGGG 180  
 Db 186 AATGTACTATTGGGGTGAAGCTTCTGAAACACTGACCTTCAAGATGACTGGAAGGGG 245  
 QY 181 GAGCTGCTCAAGACAGATCAACTCATCTGTCTGTCTGTGAGACAGAGAGAGCTTT 240  
 Db 246 GAGCTGCTCAAGACAGATCAACTCATCTGTCTGTCTGTGAGACAGAGAGAGCTTT 305  
 QY 241 GAAAAAGGCTTTTAAAGACACTTACTTCCATCATCTTGAAGAGTGAATGAG 300  
 Db 306 GAAAAAGGCTTTTAAAGACACTTACTTCCATCATCTTGAAGAGTGAATGAG 365  
 QY 301 ATTATGAGCTAGCTGTAAACCGAGCTAACCCAGATGAGATTATTTCTTAATAGTACC 360  
 Db 366 ATTATGAGCTAGCTGTAAACCGAGCTAACCCAGATGAGATTATTTCTTAATAGTACC 425  
 QY 361 CGCTTATCAATTGAGACCAAGAGATATCTGTCTTCAATCAACAGACAGGCTTATAC 420

Db 426 CGCTTATCAATTGAGACCAAGAGATATCTGTCTTCAATCAACAGACAGGCTTATAC 485  
 QY 421 AAGCCAAAGCAAGAGTGAAGTTTGGCATTTGTAACCTTCTGATTTTAAGCTTAC 480  
 Db 486 AAGCCAAAGCAAGAGTGAAGTTTGGCATTTGTAACCTTCTGATTTTAAGCTTAC 545  
 QY 481 AAAACCTCTTTAAACATCTCAATTAAGACCCCAATCAAAATTTGATCCAGTGTG 540  
 Db 546 AAAACCTCTTTAAACATCTCAATTAAGACCCCAATCAAAATTTGATCCAGTGTG 605  
 QY 541 TCACAAACAAAGTATCTTGAAGTCAATTCACAAACCTTTGACCTATCTTCCATCAATA 600  
 Db 606 TCACAAACAAAGTATCTTGAAGTCAATTCACAAACCTTTGACCTATCTTCCATCAATA 665  
 QY 601 CTTGATGATGGTCAATTCAGTTCAAGTGAATGACAGACATATTAATCAATTTTACG 660  
 Db 666 CTTGATGATGGTCAATTCAGTTCAAGTGAATGACAGACATATTAATCAATTTTACG 725  
 QY 661 GTTTCAGAAATATGTATTCACAAATTTGAAGTACTTTCAGACACCATTAATTTGTTCT 720  
 Db 726 GTTTCAGAAATATGTATTCACAAATTTGAAGTACTTTCAGACACCATTAATTTGTTCT 785  
 QY 721 ATGAATTTAGCATTTAAATGTACATACAGCAAGATATCATATGGAAGCCAGTG 780  
 Db 786 ATGAATTTAGCATTTAAATGTATTCATACAGCAAGATATCATATGGAAGCCAGTG 845  
 QY 781 AAAGAGAGCTAGAGCTTACATTTTACCTTATCTTGGGGAAGAGAAATATTT 840  
 Db 846 AAAGAGAGCTAGAGCTTACATTTTACCTTATCTTGGGGAAGAGAAATATTT 905  
 QY 841 ACAAACCACTTTAAGATTAATGATCTGCAAACTTCTCTTTAATGATGAAGATGAAA 900  
 Db 906 ACAAACCACTTTAAGATTAATGATCTGCAAACTTCTCTTTAATGATGAAGATGAAA 965  
 QY 901 AATGTATGATTTCTCAATGAGACTTTCGAAATCTGTGATCTATCTTCCCTGAGCA 960  
 Db 966 AATGTATGATTTCTCAATGAGACTTTCGAAATCTGTGATCTATCTTCCCTGAGCA 1025  
 QY 961 GTAGAAATTTTAAACAGATGACAGAAATCAGTTAAGATTTCAAGAAATTAAGCACT 1020  
 Db 1026 GTAGAAATTTTAAACAGATGACAGAAATCAGTTAAGATTTCAAGAAATTAAGCACT 1085  
 QY 1021 AATGTGTTCTTCAACCAATGATTAATGATCTGCAAACTTCTCTTTAATGATGAAGATGAAA 1080  
 Db 1086 AATGTGTTCTTCAACCAATGATTAATGATCTGCAAACTTCTCTTTAATGATGAAGATGAAA 1145  
 QY 1081 AAGCCATCTTCAACTTCAACAGCACTGTGAAGGTATCTGTGATGAGCAACCACTG 1140  
 Db 1146 AAGCCATCTTCAACTTCAACAGCACTGTGAAGGTATCTGTGATGAGCAACCACTG 1205  
 QY 1141 ACTCTTGAAGAAAGAAATATATGTATCTATACAGTGAACAGAGAACTATATCTGAG 1200  
 Db 1206 ACTCTTGAAGAAAGAAATATATGTATCTATACAGTGAACAGAGAACTATATCTGAG 1265  
 QY 1201 TACTGAGAGGATCTTAACAGTGAAGAAATGAGAAATGAGAGCTTCAAGAAATTAAT 1260  
 Db 1266 TACTGAGAGGATCTTAACAGTGAAGAAATGAGAAATGAGAGCTTCAAGAAATTAAT 1325  
 QY 1261 ACTGTCCCCCAAGTGAACCTTTTAAAGTGAATTTCCCATCTGAGAGATTCAGTGA 1320  
 Db 1326 ACTGTCCCCCAAGTGAACCTTTTAAAGTGAATTTCCCATCTGAGAGATTCAGTGA 1385  
 QY 1321 CTACAGTTGAAGGCTTATTTCTTGTGATGAAGATGACATGAGCTTCAATGTCTGTT 1380  
 Db 1386 CTACAGTTGAAGGCTTATTTCTTGTGATGAAGATGACATGAGCTTCAATGTCTGTT 1445  
 QY 1381 AAGTCTCTAGTAAGACATCAATCCAAATTAAGATGAAGATGAATTAATTAAGTGGGA 1440  
 Db 1446 AAGTCTCTAGTAAGACATCAATCCAAATTAAGATGAAGATGAATTAATTAAGTGGGA 1505  
 QY 1441 TCGCTTTTGAAGTGTGTGTAGTGGCAACAAACGATTTGAAGAGTATATGTGTA 1500  
 Db 1506 TCGCTTTTGAAGTGTGTGTAGTGGCAACAAACGATTTGAAGAGTATATGTGTA 1565

QY 1501 GTATCCAGGGGACAGTTGGTGGCTGTAGAGAAAACAAATTCACATGTTCTCTTTACA 1560  
| | | | |  
Db 1566 GTATCCAGGGGACAGTTGGTGGCTGTAGAGAAAACAAATTCACATGTTCTCTTTACA 1625  
| | | | |  
QY 1561 CCAGAAATTTCTGGACCTCCAAAAGCCCTGTATTTGTATTTATTTAGAGATGATGGG 1620  
| | | | |  
Db 1626 CCAGAAATTTCTGGACCTCCAAAAGCCCTGTATTTGTATTTATTTAGAGATGATGGG 1685  
| | | | |  
QY 1621 GAAATTTATAGAGATGTTCTTAAAAATTCCTGTTCAGCTGTGTTTTTAAAAATTAAGTAAAG 1680  
| | | | |  
Db 1686 GAAATTTATAGAGATGTTCTTAAAAATTCCTGTTCAGCTGTGTTTTTAAAAATTAAGTAAAG 1745  
| | | | |  
QY 1681 CTATTTAGAGTAAAGTAAAGTAAAGCTGACATCTGAGAAAGTCTCTTACAGATCTGTGTG 1740  
| | | | |  
Db 1746 CTATTTAGAGTAAAGTAAAGTAAAGCTGACATCTGAGAAAGTCTCTTACAGATCTGTGTG 1805  
| | | | |  
QY 1741 ACACAGCTGACCTCATAGTTGGGATTTAGCTGTTCACAAAAGTGTGATCTGATGAAT 1800  
| | | | |  
Db 1806 ACACAGCTGACCTCATAGTTGGGATTTAGCTGTTCACAAAAGTGTGATCTGATGAAT 1865  
| | | | |  
QY 1801 GCCTCTATATGATATTAACAATGGAATGTCATGTCATGTTGGAATTTATTAACAACAGGA 1860  
| | | | |  
Db 1866 GCCTCTATATGATATTAACAATGGAATGTCATGTCATGTTGGAATTTATTAACAACAGGA 1925  
| | | | |  
QY 1861 TATTTATAGGCAATGTTGATGAATTTCTTTGCACTTTTCAGAAATGTGACCTGGGTA 1920  
| | | | |  
Db 1926 TATTTATAGGCAATGTTGATGAATTTCTTTGCACTTTTCAGAAATGTGACCTGGGTA 1985  
| | | | |  
QY 1921 TTGACAGATGCAAACTTCACAGAGATTAATTTAGTGTGTTATGACAAAT 1971  
| | | | |  
Db 1986 TTGACAGATGCAAACTTCACAGAGATTAATTTAGTGTGTTATGACAAAT 2036  
| | | | |

RESULT 3  
US-10-133-013-223  
; Sequence 223, Application US/10133013  
; Publication No. US20030166903A1  
; GENERAL INFORMATION:  
; APPLICANT: Astromoff, Anna  
; APPLICANT: Bandman, Olga  
; APPLICANT: Cocks, Benjamin G.  
; TITLE OF INVENTION: GENES ASSOCIATED WITH VASCULAR DISEASE  
; FILE REFERENCE: PA-0049 US  
; CURRENT APPLICATION NUMBER: US/10/133,013  
; CURRENT FILING DATE: 2002-04-25  
; PRIOR APPLICATION NUMBER: 60/287,067  
; PRIOR FILING DATE: 2001-04-27  
; NUMBER OF SEQ ID NOS: 271  
; SOFTWARE: PERL Program  
; SEQ ID NO 223  
; LENGTH: 1300  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20030166903A1 3658034CB1  
US-10-133-013-223

Query Match 17.0%; Score 730.8; DB 13; Length 1300;  
Best Local Similarity 93.6%; Pred. No. 4.8e-179;  
Matches 805; Conservative 0; Mismatches 2; Indels 53; Gaps 2;  
QY 3479 AGGNAATCCCAATTTATGAGTGGCTTAAGCAGGCAAAAGAAATAGCTTGGGTGTTTGCAT 3538  
| | | | |  
Db 1 AGGNAATCCCAATTTATGAGTGGCTTAAGCAGGCAAAAGAAATAGCTTGGGTGTTTGCAT 60  
| | | | |  
QY 3539 CTACTCAGATATACCACTGTGGCTTTAAAGGCTGTCTGATTTGACAGCCCTTAATGAATA 3598  
| | | | |  
Db 61 CTACTCAGATATACCACTGTGGCTTTAAAGG--CTGTCTGATTTTGCAGCCCTTAATGAATA 118  
| | | | |  
QY 3599 CAGAAAGCAAAATATCCAACTGACCGTGAAGGGGCTTACGCAAGTCTCT----- 3651  
| | | | |

Db 119 CAGAAAGCAAAATATCCAACTGACCGTGAAGGGGCTTACGCAAGTCTCTGAAGT 178  
| | | | |  
QY 3652 -----CTGCTGTGTACAGC 3667  
| | | | |  
Db 179 TTCTGATTTGACACACACACACCGCTTACTCTTTCAGACAGAGACCTTGTGTGTACAGC 238  
| | | | |  
QY 3668 CAATGCGATTAATATTTCCGCAAAATGTTTGTATTTGTCTATTTGTCACTCAATGTTG 3127  
| | | | |  
Db 239 CAACGCGATTAATATTTCCGCAAAATGTTTGTATTTGTCTATTTGTCACTCAATGTTG 298  
| | | | |  
QY 3728 TATTTATATGAGAGCTTCTGGGCTTCTTGAAGACGAATATCTATCCAAATATCAAGAG 3787  
| | | | |  
Db 299 TATTTATATGAGAGCTTCTGGGCTTCTTGAAGACGAATATCTATCCAAATATCAAGAG 358  
| | | | |  
QY 3788 CTTTGAATTTAGATGTTGCTGTAAAAAGAAATATTAAGTATCTCAATCATGTGGATTTGA 3847  
| | | | |  
Db 359 CTTTGAATTTAGATGTTGCTGTAAAAAGAAATATTAAGTATCTCAATCATGTGGATTTGA 418  
| | | | |  
QY 3848 ATGTGTACAACTTTTGGGCCCCGGTGAAGTGCATGCTCTTATGAAAGTTAAAC 3907  
| | | | |  
Db 419 ATGTGTACAACTTTTGGGCCCCGGTGAAGTGCATGCTCTTATGAAAGTTAAAC 478  
| | | | |  
QY 3908 TATTTAGTGGCTTTATGATGTCCTTTCAGAGCAATTTCTGACGAGACAGTGAAGAG 3967  
| | | | |  
Db 479 TATTTAGTGGCTTTATGATGTCCTTTCAGAGCAATTTCTGACGAGACAGTGAAGAG 538  
| | | | |  
QY 3968 TGGAAATATGATCATGAGAACTCAACCTTATTTAGATTTCTGTAATGAAACCAAGTTT 4027  
| | | | |  
Db 539 TGGAAATATGATCATGAGAACTCAACCTTATTTAGATTTCTGTAATGAAACCAAGTTT 598  
| | | | |  
QY 4028 GTGTTAATATTTCTGCTGTGAGAACTTTAAAGTTTCAATTAACCAAGATGCTTCAAGTGT 4087  
| | | | |  
Db 599 GTGTTAATATTTCTGCTGTGAGAACTTTAAAGTTTCAATTAACCAAGATGCTTCAAGTGT 658  
| | | | |  
QY 4088 CCATAGTGCATTTCTATGAGCCCAAGAGACAGCCGCTGAGAACTTCAACTCTGAAGTGA 4147  
| | | | |  
Db 659 CCATAGTGCATTTCTATGAGCCCAAGAGACAGCCGCTGAGAACTTCAACTCTGAAGTGA 718  
| | | | |  
QY 4148 AGCTGTCTCTGCTGAGCTTTGACAGTATGTCAGAGGCTCCGCTCTTGTGAGAGATGAG 4207  
| | | | |  
Db 719 AGCTGTCTCTGCTGAGCTTTGACAGTATGTCAGAGGCTCCGCTCTTGTGAGAGATGAG 778  
| | | | |  
QY 4208 CTTGAGGCTCCATCATCATCTCTTCAGTCAATTTTATTTTCTGTTCAAGCTTCTGACT 4267  
| | | | |  
Db 779 CTTGAGGCTCCATCATCATCTCTTCAGTCAATTTTATTTTCTGTTCAAGCTTCTGACT 838  
| | | | |  
QY 4268 TTATGAACTTTGGCTGTGA 4287  
| | | | |  
Db 839 TTATGAACTTTGGCTGTGA 858  
| | | | |

RESULT 4  
US-10-160-162-48  
; Sequence 48, Application US/10160162  
; Publication No. US20030166541A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: 83 Human Secreted Proteins  
; FILE REFERENCE: P2012P2  
; CURRENT APPLICATION NUMBER: US/10/160,162  
; CURRENT FILING DATE: 2002-06-04  
; PRIOR APPLICATION NUMBER: 60/295,558  
; PRIOR FILING DATE: 2001-06-05  
; PRIOR APPLICATION NUMBER: 09/236,557  
; PRIOR FILING DATE: 1999-01-26  
; PRIOR APPLICATION NUMBER: PCT/US98/15949  
; PRIOR FILING DATE: 1998-07-29  
; PRIOR APPLICATION NUMBER: 60/054,212  
; PRIOR FILING DATE: 1997-07-30  
; PRIOR APPLICATION NUMBER: 60/054,209  
; PRIOR FILING DATE: 1997-07-30  
; PRIOR APPLICATION NUMBER: 60/054,234  
; PRIOR FILING DATE: 1997-07-30

```

? PRIOR APPLICATION NUMBER: 60/054,218
? PRIOR FILING DATE: 1997-07-30
? PRIOR APPLICATION NUMBER: 60/054,214
? PRIOR FILING DATE: 1997-07-30
? PRIOR APPLICATION NUMBER: 60/054,236
? PRIOR FILING DATE: 1997-07-30
? PRIOR APPLICATION NUMBER: 60/054,215
? PRIOR FILING DATE: 1997-07-30
? PRIOR APPLICATION NUMBER: 60/054,211
? PRIOR FILING DATE: 1997-07-30
? PRIOR APPLICATION NUMBER: 60/054,217
? PRIOR FILING DATE: 1997-07-30
? PRIOR APPLICATION NUMBER: 60/054,213
? PRIOR FILING DATE: 1997-07-30
? PRIOR APPLICATION NUMBER: 60/055,968
? PRIOR FILING DATE: 1997-08-18
? PRIOR APPLICATION NUMBER: 60/055,969
? PRIOR FILING DATE: 1997-08-18
? PRIOR APPLICATION NUMBER: 60/055,972
? PRIOR FILING DATE: 1997-08-18
? PRIOR APPLICATION NUMBER: 60/056,561
? PRIOR FILING DATE: 1997-08-19
? PRIOR APPLICATION NUMBER: 60/056,534
? PRIOR FILING DATE: 1997-08-19
? PRIOR APPLICATION NUMBER: 60/056,729
? PRIOR FILING DATE: 1997-08-19
? PRIOR APPLICATION NUMBER: 60/056,543
? PRIOR FILING DATE: 1997-08-19
? PRIOR APPLICATION NUMBER: 60/056,727
? PRIOR FILING DATE: 1997-08-19
? PRIOR APPLICATION NUMBER: 60/056,554
? PRIOR FILING DATE: 1997-08-19
? PRIOR APPLICATION NUMBER: 60/056,730
? PRIOR FILING DATE: 1997-08-19
? NUMBER OF SEQ ID NOS: 353
? SOFTWARE: Patentin Ver. 2.0
? SEQ ID NO 48
? LENGTH: 875
? TYPE: DNA
? ORGANISM: Homo sapiens
US-10-160-162-48

```

Query Match	Similarity	12.6%	Score 541.6	DB 13	Length 875
Best Local	Similarity	99.8%	Pred. No. 6.4e-130		
Matches	541	Conservative	1	Mismatches	0
				Indels	0
				Gaps	0
QY	3746	CTGGGCTCTTGAAGACGAGAGATCTATCCAAATCAAGAGCCCTTGATTAGATGTG	3805		
Db	15	CTGGGCTCTTGAAGACGAGAGATCTATCCAAATCAAGAGCCCTTGATTAGATGTG	74		
QY	3806	CTGTAAAGAAATPAAAGATGATCTCAATCATGTGAGTTTGAATGTGTGACAGCTTTT	3865		
Db	75	CTGTAAAGAAATPAAAGATGATCTCAATCATGTGAGTTTGAATGTGTGACAGCTTTT	134		
QY	3866	CGGGCCCGGATGAGAGTGCATGGCTCTTATGGAAGTTAACTCATTAAGTGGCTTTATGG	3925		
Db	135	CGGGCCCGGATGAGAGTGCATGGCTCTTATGGAAGTTAACTCATTAAGTGGCTTTATGG	194		
QY	3926	TGCCTTCGAAAGCAATTTCTCGACGAGACAGTGAAGAAAGTGAAATATGATCATGAA	3985		
Db	195	TGCCTTCGAAAGCAATTTCTCGACGAGACAGTGAAGAAAGTGAAATATGATCATGAA	254		
QY	3986	AATCTAACCTCTATTTAATATTCGTAAATGAAACCCAGTTTGTGTAAATATTCCTGCTG	4045		
Db	255	AATCTAACCTCTATTTAATATTCGTAAATGAAACCCAGTTTGTGTAAATATTCCTGCTG	314		
QY	4046	TGAGAAATTTAAAGTTTCAAAATACCAAGATGCTTCAGTGTCCATAGTGATTTACTAAG	4105		
Db	315	TGAGAAATTTAAAGTTTCAAAATACCAAGATGCTTCAGTGTCCATAGTGATTTACTAAG	374		
QY	4106	AGCCAAGAGACAGCGGCTGAGAAAGTTACAACTGTGAAGTGAAGCTGTCTCTGTGACC	4165		
Db	375	AGCCAAGAGACAGCGGCTGAGAAAGTTACAACTGTGAAGTGAAGCTGTCTCTGTGACC	434		

```

OY      4166  TTTCAGATGATCCAGGGCTGCCGCTCTTGATGAGATGAGAGCTTCAGGGCTCCCATCATC  4225
Db      435   TTTCAGGATGATCCAGGGCTGCCGCTCTTGATGAGATGAGAGCTTCAGGGCTCCCATCATC  494
OY      4226  ACTCTTCAGTCATTTTATTTCTGTTCAAGCTTCGTACTTATATGAACTTTGGCTGT  4285
Db      495   AMCTTCAGTCATTTTATTTCTGTTCAAGCTTCGTACTTATATGAACTTTGGCTGT  554
OY      4286  GA  4287
Db      555  GA  556

```

```

US-09-820-649-48
RESULT 5
; Sequence 48, Application US/09820649
; Publication No. US20030199683A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 83 Human Secreted Proteins
; FILE REFERENCE: P2012P1
; CURRENT APPLICATION NUMBER: US/09/820,649
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US/09/236,557
; PRIOR FILING DATE: 1999-01-26
; PRIOR APPLICATION NUMBER: PCT/US98/15949
; PRIOR FILING DATE: 1998-07-29
; PRIOR APPLICATION NUMBER: 60/054,212
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/054,209
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/054,214
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/054,218
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/054,214
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/054,216
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/054,215
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/054,211
; PRIOR FILING DATE: 1997-07-30
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 353
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 48
; LENGTH: 875
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-820-649-48

```

Query Match	12.6%	Score 541.6	DB 13	Length 875
Best Local Similarity	99.8%	Pred. No. 6.4e-130		
Matches 541	Conservative 1	Mismatches 0	Indels 0	Gaps 0
QY	3746	CTGGGCTCTTCTAGAAGC	GAAGATCTATCCAAATCA	GAAGACCTTTGATTAGATGTG 3805
Db	15	CTGGGCTCTTCTAGAAGC	GAAGATCTATCCAAATCA	GAAGACCTTTGATTAGATGTG 74
QY	3806	CTGTAAAAAGAAAATAA	AGATGATTCATATCTGAT	TGGAATGCTGTATCAAGCTTTT 3865
Db	75	CTGTAAAAAGAAAATAA	AGATGATTCATATCTGAT	TGGAATGCTGTATCAAGCTTTT 134
QY	3866	CGGGCCCCGGGTAGAG	TGCGCATGGCTTTATGAA	GTAACTTTAATGCGCTTTATG 3925
Db	135	CGGGCCCCGGGTAGAG	TGCGCATGGCTTTATGAA	GTAACTTTAATGCGCTTTATG 194
QY	3926	TGCTCTCAGAAACCA	TTTCTCTGTAGCGAC	ACGTGAAGAAAGTGGAAATGATCATGGA 3985
Db	195	TGCTCTCAGAAACCA	TTTCTCTGTAGCGAC	ACGTGAAGAAAGTGGAAATGATCATGGA 254

Qy 3986 AACTCAACCTCTATTAGATTCTGTAATAAACCAGTTTGTGTAAATATTCCTGCTG 4045  
Db 255 AACTCAACCTCTATTAGATTCTGTAATAAACCAGTTTGTGTAAATATTCCTGCTG 314  
Qy 4046 TGAGAACTTTAAAGTTTCAAAATCCCAAGATGCTGTGCTCAATGATGATTAATG 4105  
Db 315 TGAGAACTTTAAAGTTTCAAAATCCCAAGATGCTGTGCTCAATGATGATTAATG 374  
Qy 4106 AGCCAAAGAGACAGCGGTGAGAAAGTTCAACTGTAAGTAAGCTGCTCTGTGACC 4165  
Db 375 AGCCAAAGAGACAGCGGTGAGAAAGTTCAACTGTAAGTAAGCTGCTCTGTGACC 434  
Qy 4166 TTTGAGATGATGTCAGGAGCTGCTGCTGTTGAGATGAGCTTCAGGCTCCATCATC 4225  
Db 435 TTTGAGATGATGTCAGGAGCTGCTGCTGTTGAGATGAGCTTCAGGCTCCATCATC 494  
Qy 4226 ACTCTGAGATTTTATTTTCTGTTCAAGCTTCTGATTAATGAACTTTGGCTGT 4285  
Db 495 ACTCTGAGATTTTATTTTCTGTTCAAGCTTCTGATTAATGAACTTTGGCTGT 554  
Qy 4286 GA 4287  
Db 555 GA 556

## RESULT 6

US-09-960-352-12867  
; Sequence 12867, Application US/09960352  
; Patent No. US20020137139A1  
; GENERAL INFORMATION:  
; APPLICANT: Warren, Wesley C.  
; APPLICANT: Tao, Mengling  
; APPLICANT: Byatt, John C.  
; APPLICANT: Mathalagan, Nagappan  
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND  
; FILE REFERENCE: 16511.006/37-21(10298)C  
; CURRENT APPLICATION NUMBER: US/09/960.352  
; CURRENT FILING DATE: 2001-09-24  
; NUMBER OF SEQ ID NOS: 15112  
; SEQ ID NO 12867  
; LENGTH: 354  
; TYPE: DNA  
; ORGANISM: Bos taurus  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (285)  
; OTHER INFORMATION: unsure at all n locations  
; OTHER INFORMATION: Clone ID: 55-LIH3057-014-01-K1-F12  
US-09-960-352-12867

Query Match 6.3%; Score 272; DB 10; Length 354;

Best Local Similarity 89.1%; Pred. No. 5, 1e-60; Matches 293; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

Qy 2084 TTTGCTAGACACCAATGAGTTTACAGATTTCACGAATTTGAAGTAATCTGACTG 2143  
Db 26 TTTGCTAGACACCAATGAGTTTACAGATTTCACGAATTTGAAGTAATCTGACTG 85  
Qy 2144 ATTCTATCATCTTTGGGTGCTACTGTTTGTGATCTTGAGAGACTGGGCTTGAC 2203  
Db 86 ATTCTATCATCTTTGGGTGCTACTGTTTGTGATCTTGAGAGACTGGGCTTGAC 145  
Qy 2204 TAAACAATACCTCCAGTGAAGCTCCAAAGCTTCCAAACATTTTCAATTTTGAATCTC 2263  
Db 146 TAAACAATACCTCCAGTGAAGCTCCAAAGCTTCCAAACATTTTCAATTTTGAATCTC 205  
Qy 2264 CCTACTCTGATACAGAGTGAAGATTTGCTTGAATAATCAATATTAATTTGA 2323  
Db 206 CATCTCTGATACAGAGTGAAGATTTGCTTGAATAATCAATATTAATTTGA 265  
Qy 2324 AAGATGCCACTGAGGTTAAGTAATCATTTGAGAAAAGTGACAAATTTGATATTGAATGA 2383  
Db 2383 AAGATGCCACTGAGGTTAAGTAATCATTTGAGAAAAGTGACAAATTTGATATTGAATGA 2383

Db 266 AAGATGCCACTGAGGTTAAGTAATCATTTGAGAAAAGTGATGCTTTGATATTGAATG 325  
Qy 2384 CTTCAGTGAATAATTAATGCCAGAGCCAC 2412  
Db 326 CTTCAGTGAATAATTAATTAATGCCAGAGCCAC 354

## RESULT 7

US-10-027-632-204326/c  
; Sequence 204326, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027.632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 204326  
; LENGTH: 689  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-204326

Query Match 4.5%; Score 194.4; DB 13; Length 689;

Best Local Similarity 99.5%; Pred. No. 1, 3e-39; Matches 195; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3352 GAAGTGGCATGGAATTTGGGTGCTATCAGATGCAACTTTCTGACTCTGGACCCA 3411  
Db 485 GAAGTGGCATGGAATTTGGGTGCTATCAGATGCAACTTTCTGACTCTGGACCCA 426  
Qy 3412 CGCTCCCTGATATTGAAGTTGACGCTATGCACTGCTTCACACTTTTACAATTTGAG 3471  
Db 425 CGCTCCCTGATATTGAAGTTGACGCTATGCACTGCTTCACACTTTTACAATTTGAG 366  
Qy 3472 ACTTTGAGGAATTCGAATTTAGAGTGGCTTAAGCAGCAAGAAATGCTGGGTGT 3531  
Db 365 ACTTTGAGGAATTCGAATTTAGAGTGGCTTAAGCAGCAAGAAATGCTGGGTGT 306  
Qy 3532 TTTGATCTACTCAGG 3547  
Db 305 TTTGATCTACTCAGG 290

## RESULT 8

US-10-027-632-204326/c  
; Sequence 204326, Application US/10027632  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027.632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12

```

1 PRIOR APPLICATION NUMBER: US 60/198,676
2 PRIOR FILING DATE: 2000-04-20
3 PRIOR APPLICATION NUMBER: US 60/193,483
4 PRIOR FILING DATE: 2000-03-29
5 PRIOR APPLICATION NUMBER: US 60/185,218
6 PRIOR FILING DATE: 2000-02-24
7 PRIOR APPLICATION NUMBER: US 60/167,363
8 PRIOR FILING DATE: 1999-11-23
9 PRIOR APPLICATION NUMBER: US 60/156,358
10 PRIOR FILING DATE: 1999-09-28
11 PRIOR APPLICATION NUMBER: US 60/146,002
12 PRIOR FILING DATE: 1999-08-09
13 NUMBER OF SEQ ID NOS: 325720
14 SOFTWARE: FastSeq for Windows Version 4.0
15 SEQ ID NO 204326
16 LENGTH: 689
17 TYPE: DNA
18 ORGANISM: Human
19 US-10-027-652-204326

```

Query Match	4.5%	Score 194.4;	DB 14;	Length 689;
Best Local Similarity	99.5%;	Pred. No. 1.3e-39;		
Matches 195; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0

QY	3352	GAAAGTGCATGCATTTCTGGGTGTCACTACAGATCCAACTTTGCACTCTGGGACCA	3411
QY	3352	GAAAGTGCATGCATTTCTGGGTGTCACTACAGATCCAACTTTGCACTCTGGGACCA	3411
Db	485	GCAAGTGCATGCATTTCTGGGTGTCACTACAGATCCAACTTTGCACTCTGGGACCA	426
QY	3412	CGCTCCCTGATATTGAAGTTGCAGCCTATGCACTGCTCTCACTTCTTACAAATTTCAG	3477
Db	425	CGCTCCCTGATATTGAAGTTGCAGCCTATGCACTGCTCTCACTTCTTACAAATTTCAG	366
QY	3472	ACTTCTGAGGGAATTCCAATTATGAGGTGGCTTAAGCAGCAAGAAATATGCTTGGTGT	3533
Db	365	ACTTCTGAGGGAATTCCAATTATGAGGTGGCTTAAGCAGCAAGAAATATGCTTGGTGT	306
QY	3532	TTTGATCTACTCAGG	3547
Db	305	TTTGATCTACTCAGG	290

```

RESULT 9
US-10-316-253-266
: Sequence 266, Application US/10316253
: Publication No. US20030162706A1
: GENERAL INFORMATION:
: APPLICANT: The Procter & Gamble Company
: APPLICANT: Peters, Kevin
: APPLICANT: Thompson, Larry
: APPLICANT: Wang, Feng
: APPLICANT: Greis, Kenneth
: TITLE OF INVENTION: Angiogenesis Modulating Proteins
: FILE REFERENCE: 8865M
: CURRENT APPLICATION NUMBER: US-10/316,253
: CURRENT FILING DATE: 2002-12-10
: PRIOR APPLICATION NUMBER: US 60/355,295
: PRIOR FILING DATE: 2002-02-08
: NUMBER OF SEQ ID NOS: 308
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 266
: LENGTH: 4677
: TYPE: DNA
: ORGANISM: Rattus norvegicus
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (52)..(4554)
: OTHER INFORMATION:
: US-10-316-253-266

```

Query Match	3.0%;	Score 130;	DB 13;	Length 4677;
Best Local Similarity	47.3%;	Pred. No. 3e-22;		
Matches 575;	Conservative	0;	Mismatches 550;	Indels 51;
			Gaps	4

OY	2056	GTCCGGAAGA	CAATTTTCCAGAGAC	CTTGGATTTGGCTAGAC	CAACCAAC	AGGGTTACAGATT	2115
Db	2284	GTGGGGAAGT	ACTTCCCGGAAC	CTGGATCTGGGACAT	TGGTCCGCTGGACTTAT	CCGGT	2343
OY	2116	TACCAAGAA	TTTGAAGTAACTGTA	ACCTGAATCTATCA	CTTCTTGGTGGCTACTGATTT		2175
Db	2344	GACGGTAGT	TGCGGGTAAGGTC	CTCGACACCATCA	CTGATGTGAAGGCCAGTGCA	TTTC	2403
OY	2176	GTAATCTCTG	AGGACCTGGGCTTT	GACTAACTACTCA	GATGGAGCTTCAAGCTTTC		2235
Db	2404	TGCTTGCTG	GAACCTAC	TGGCCTTGGCTCTCTCC	CAACA---TCTCCATTAAGTCTTC		2460
OY	2236	CAACCA	TTTTTCAATTTTTTGAAT	CTTCCCTACCTCTGTAT	TCAGAGGGTGAAGAA	TTTGGCT	2295
Db	2461	CAGCCCTTCT	CTCGAGCTCACTCTCC	CTACTCTGTGGTTG	AGGGCAAGCA	TTTATTC	2520
OY	2296	TTGGAATA	TTAACTATTTCAATTTAT	TGGAAGATGCCA	CTGAGGTTAAGTAA	TATCATTTAG	2355
Db	2521	CTCAAGCT	CTCCGACTCACTAC	TGCTTCACTGCA	ATTGGATTCATGTGAG	CTTAGG	2580
OY	2356	AAA-----					2385
Db	2581	ATGCTCTCTG	ATTTCTCGAGTCC	CTGTGGGGAGCAT	GAAGATTTCTCATTTGATCTGT		2640
OY	2386	TCAGTGAAT	TAATAATGCGACAGG	CAACACAC	CCCTTGTGTTCCCACTGAGAT	GGG	2445
Db	2641	GGAATATA	AAAGAAAMAC	CGTGCTCTGGGCTGTG	ACCCCAAGTGC	GTGGGGAGGTGAC	2700
OY	2446	GCAACTTT	CTTTTCCCATCAGG	CCAAACATCTG	GGAAATAATTC	CTATCACAGTACA	2505
Db	2701	TTTCACTACT	GTGAGAACTTTTGCA	GTCTTCAGAA	CTGTGTGGCA	TAAAGTGGCAGAA	2760
OY	2506	GCTCTTTC	ACCCCACTGCTTGT	GATGCTGTCA	CCCAATGATTTT	AGTAAAGCTGAAGA	2565
Db	2761	GTGCGAC	CCCTTGTA	CAGAGGACACTGT	ATGTGAAGCCCGTA	ATATGTGAGCTGAAGA	2820
OY	2566	ATGAAAAA	ATCATTTTCA	CAATCCATCTTAT	TAGACTTGAC	TAAATAGCTACAGT	2625
Db	2821	ATTGAGA	AGGAACAAAGT	CAACACACTGT	TATGCCCCA	CAGATGTCTGATTAACA----	2876
OY	2626	ACCCTGAAA	CTTTGATTTCT	CAATTTCC	CTTAATCAGTGA	CTGGCAGTGAAGA	2685
Db	2877	-----					2931
OY	2686	CAGATCA	CTGCAATTTGAGAT	TGTTCTTGTCTT	CCATCATATG	GGCTTAGGCTCAATGAT	2745
Db	2932	ACA	CAGTCCGTTTTGGG	CGATATATCTAG	GGCTGCGCATG	CAAAACCTCCAGAA	2991
OY	2746	CGGATG	CTTATGCTGTGT	GTAACAGACAT	GTATTAATTTT	GCTCCAAATATTTACAT	2805
Db	2992	CAGATG	CCCTATGTGCTGTGGG	GAACAAACAT	GGTCTCTTTC	CGCTTAACATCTACGTT	3051
OY	2806	TTGATAT	CTGACTAATAAAGAA	CAACATCA	CGATTAATT	TGAAAGAAAGAAAGCTCTTCA	2865
Db	3052	CTGGA	GTATCTCATGTAG	ACACAGAC	CTCAAGGGCGAT	CAAGTCCAAAGCCATTAAC	3111
OY	2866	TTTTATG	AGCAAGGTTAC	CAGAGGAAC	CTTCTATCAG	AGGGAAGATGGCTTTTCAGT	2925
Db	3112	TACCTCAT	CAAGCGGATCC	AGAAGCAGCTGA	CTATACGACACAGT	ACGTTCAITACACG	3171
OY	2926	GCTTTGGG	AATTAATG-----				2976
Db	3172	ACATTT	CGGGGACCGCGGTAT	AGGCAAGTCA	GGGAAACCTTGGCTCA	CTGCACTTGTGTG	3231
OY	2977	TTTAAGAT	GTTTCTTGAAG	CCGATCTTATCAT	ATAGATTTT	GATCAGATGTGTTACACAGA	3036
Db	3232	CTCAAG	CCCTTGGCTCAAG	CTCAAGTCAAT	CTATATATTA	AAAAAGACACATCAACAAT	3291
OY	3037	ACATAC	CACTTGGCTTAAG	ACATCAAGAA	ATCAACG	GTAAATTTTGGATTCAGGAAGA	3096
Db	3292	GCTTTCA	ATTGGCTCTCGAT	GAACAAAGG	GAACG	GTGTTTCCAAAGTCTGATCC	3351
OY	3097	GTAATCAT	ATGTAGCTTTCA	AGGTGGCA	TATAAAGTCA	GTAAACATTTACAGCCATATAT	3156

Db 3352 CTGCTCAACAATGCGATGAAAGGTGGTGTGATGATGAAAGTGAACACTCTCTGCTATATC 3411  
Qy 3157 GTAACCTCTCTCCGGGATATAGAAAGTATACAGCCTAACATGATGAGCAGAGCTATC 3216  
Db 3412 ACCATTGCTCTGCTGGAGATGCCCTGCTGCTACACAGTGTGTTGTAAGTCTTA 3471  
Qy 3217 CATTTTGGAGTCTG 3232  
Db 3472 TTCTGCTGGAACGG 3487

RESULT 10  
US-10-316-253-268  
; Sequence 268, Application US/10316253  
; Publication No. US20030162706A1  
; GENERAL INFORMATION:  
; APPLICANT: The Procter & Gamble Company  
; APPLICANT: Peters, Kevin  
; APPLICANT: Thompson, Larry  
; APPLICANT: Wang, Feng  
; APPLICANT: Greis, Kenneth  
; TITLE OF INVENTION: Angiogenesis Modulating Proteins  
; FILE REFERENCE: 8865M  
; CURRENT APPLICATION NUMBER: US/10/316,253  
; CURRENT FILING DATE: 2002-12-10  
; PRIOR APPLICATION NUMBER: US 60/355,295  
; PRIOR FILING DATE: 2002-02-08  
; NUMBER OF SEQ ID NOS: 308  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 268  
; LENGTH: 4701  
; TYPE: DNA  
; ORGANISM: Rattus norvegicus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (95)..(4597)  
; OTHER INFORMATION:  
US-10-316-253-268

Query Match 3.0%; Score 130; DB 13; Length 4701;  
Best Local Similarity 47.3%; Pred. No. 3e-22;  
Matches 575; Conservative 0; Mismatches 590; Indels 51; Gaps 4;

Qy 2056 GTCCGAAGATTTTCCAGAGCTTGATTTGGCTAGACACCAACATGGTTACAGATT 2115  
Db 2327 GTGCGGAAGTACTTCCCGGAACCTGATCTGGACATGTCGCGCTGACTTATCCGCT 2386  
Qy 2116 TACCAAGATTTGAAGTATGATCTGATCTTATCTTCTTGGTGGCTACTGTTT 2175  
Db 2387 GACGGTGAAGTGGCGGTAAAGTCCCTGACACACATCACTGAGTGAAGCCAGTGCATT 2446  
Qy 2176 GTGATCTTGAGAGCTGGGTCTTGAGTAACTCACTCCAGTGAAGCTCCAGCCTTC 2235  
Db 2447 TGCTGTCTGGAACACTAGCGCTTGCTCTCTCCACCA---TCTCCATTAAGTCTTC 2503  
Qy 2236 CAACCATTTTTCATTTTGTGATCTTCCCTACTCTGTATCAAGGTAAGAAATTTGCT 2295  
Db 2504 CAGCCCTTCTCTGAGACTCACTTCCCTACTCTGTGTTCCAGGCGAAGCATTTATC 2563  
Qy 2296 TTGGAATATATATTTCAATTAATTTGAAGATGACCTGAGTTAAGTATATGAG 2355  
Db 2564 CTCGAAGTACCGTATCTCACTACATGCTCCTGATTCGATTCATGAGCTAGAG 2623  
Qy 2356 AAA-----AGTCAAAATTTGATATTTCTATGACT 2385  
Db 2624 ATGCTCTCTGATTTCTGCAAGTCCCTGTGGGAGCCATGAAGATTTCTATTGACATCTGT 2683  
Qy 2386 TCAAGTAAATTAATGACACAGGCGACAGACAGACCCCTTGGTTCCAGTGAAGATGG 2445  
Db 2684 GGAATTAAGAAAGAAACCGTGTCTGGGCTGTGACCCCAAAAGTCCGTGGGAGGTGAAC 2743  
Qy 2446 GCAACTGTCTTTTCCATGAGGCCAACACATCTGGAGAAATTTCTATACAGTCA 2505

Db 2744 TTACAGCTTCTGAGAGAGCTTTTGACATCTCCAGAACTGTGTGGCAATTAAGTGGCAGAA 2803  
Qy 2506 GCTCTTACCAACCTGCTTGTGATGCTGACACCAAGATGATTTTGAAGGCTGAAGGA 2565  
Db 2804 GTGCGAGCCCTTTGACAGAAAGACACTGTGTAGTGAAGCCGTAATATGTCGAGCTGAAGGA 2863  
Qy 2566 ATGAAAAATCATATTTCAATCCATCTTATTAAGACTTGACTGACATAGGCTACAGAGT 2625  
Db 2864 ATTGAGAAAGAACAAAGCTACACACACTGTATTGCCCACAAAGTGTGAGTTA----- 2919  
Qy 2626 ACCCTGAAAACCTTGAATTTCTCATTTCTCTCTAATACAGTGAAGTGAAGGTT 2685  
Db 2920 -----GGAAAAATGGAACACTGATCTTCCAGCCAAATGTGTTGAAGATCTGCGAGGCT 2974  
Qy 2686 CAGATCACTGCATTTGAGAGATGTTCTTGTGCTCTTCATCAATAGGCTTAGCCTCATGATT 2745  
Db 2975 ACACAGTCCGTTTGGGCGATATCTAGGCTCTGCGATGCAAAACCTCCAGAAATCTTCTC 3034  
Qy 2746 CGATGCTTATGCGCTGTGTGAAACAGAACATGATAATTTTGTCTCCAAATATTACATT 2805  
Db 3035 CAGATGCCATATGGCTGTGGGAAACAAACATGGTCTTCCGCTTAATCATCTACGTT 3094  
Qy 2806 TTGATATCTGACTATAAAAGAAACAACTGACATTAATTTGAAGAAAAGCTCTTTCA 2865  
Db 3095 CTGAGATATCTCAATGACACACAGCAGCTGACAGAGCGATCAAGTCCAAAGCCATTAGC 3154  
Qy 2866 TTTATGAGGAGGTTACAGAGAACTCTCTATCAGAGGGAAGATGCTCTTTAGT 2925  
Db 3155 TACTCTACAGCGGATCCAGAGAGCTGAATCATATGACACAGTGCAGTTCAATACAC 3214  
Qy 2926 GCTTTTGGAAATTAAG-----ACCTTCTGGAGCACTTGTGTGACGTTTGT 2976  
Db 3215 ACATTTGGGAGACGGCGGTATAGGACACAGTCAAGGAAACCTTGGCTCATGATTTGTG 3274  
Qy 2977 TTAAGATGTTTCTTGAAGCCGATCTTATCATGATTTGATCAGAAATGTTTACACAGA 3036  
Db 3275 CTCAGGCTTGTGCTCAAGCTCAAGTCAATCATATATTAAGAAAGACACACATACAAAT 3334  
Qy 3037 ACATACACTGGCTTAAGAGACATCCAGAAATCCAAAGGTGAATTTTGGATCCAGAGA 3096  
Db 3335 GCTTTCATTTGGCTCTCTGATGAACAAAGGAGACGTTGTTCCAACTGTGAGATCC 3394  
Qy 3097 GTGATTCATAGTGGCTTCAAGGTGCGCATTAAGTCCAGTAACTTACAGCCTATATT 3156  
Db 3395 CTGCTCAACATGAGTGAAGAGGTGTGTGATGATGAATGACACTCTGCTATATC 3454  
Qy 3157 GTAACCTCTCTGCGATATTAAGAAATACAGCTTAACATTTGATGTCAGAGTCTATC 3216  
Db 3455 ACCATGCTCTGCTGAGATGCCCTGCTGCTACCTACAGTGTGTTGTAATGCTCTA 3514  
Qy 3217 CATTTTGGAGTCTG 3232  
Db 3515 TTCTGCTGGAACGG 3530

RESULT 11  
US-09-880-107-3727  
; Sequence 3727, Application US/09880107  
; Patent No. US20020142981A1  
; GENERAL INFORMATION:  
; APPLICANT: Horne, Darci T.  
; APPLICANT: Vockley, Joseph G.  
; APPLICANT: Scherf, Uwe  
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer  
; FILE REFERENCE: 44921-5028-WO  
; CURRENT APPLICATION NUMBER: US/09/880,107  
; CURRENT FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: US 60/211,379  
; PRIOR FILING DATE: 2000-06-14  
; PRIOR APPLICATION NUMBER: US 60/237,054  
; PRIOR FILING DATE: 2000-10-02

NUMBER OF SEQ ID NOS: 3950  
SOFTWARE: Patenclin Ver. 2.1  
SEQ ID NO 3727  
LENGTH: 4615  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: Genbank Accession No. US20020142981A1 X54380  
US-09-880-107-3727

Query Match 3.0%; Score 126.8; DB 10; Length 4615;  
Best Local Similarity 51.6%; Pred. No. 2e-21;  
Matches 318; Conservative 0; Mismatches 292; Indels 6; Gaps 1;

QY 2638 TTGAGTTTCTCATTTCTCTCTAATACAGTACGTGCGAGTGAAGAGTTGACATCTGCA 2697  
DB 2838 TTGTCCTTGAAGCTCCCATCAATGTGTGCAAGATCTGCCAGAGCTTCTTTCAGTT 2897  
QY 2638 ATTGAGATGTTCTTGTGCTTCCATCAATGGCTTAGCCTCATTTGATTCGATCCTTAT 2757  
DB 2838 CTGGGTGACATATTATGTTCTCTCTATGCAAAATATACAAATCTCCTCAGATGCCATAT 2957  
QY 2758 GGCTGTGTAACAGAACATGATTAATTTGCTCCAAATATTACATTTTGGATTATCTG 2817  
DB 2958 GGCTGTGTAACAGAACATGATTAATTTGCTCCAAATATTACATTTTGGATTATCTG 3017  
QY 2818 ACTAAAGAAACAACTGACAGATTAATTGAAAGAAAGCTCTTTCATTATTAAGGCAA 2877  
DB 3018 AATGAACCCAGACAGTACGAGAGATCAAGGCCAAGCCGTTGGCTATCTCATCT 3077  
QY 2878 GGTTCACAGAGAACTTCTATACAGAGGAAGATGGCTCTTTCAGTCTTTGGG--- 2934  
DB 3078 GGTTCACAGAGAACTTCTATACAGAGGAAGATGGCTCTTTCAGTCTTTGGG--- 2934  
QY 2935 ---AATTATGACCCCTTCTGGAGCACTTGTGTGAGCTTTTGTATTAAGATTTTCTT 2991  
DB 3138 CGATATGGCAGGAACCAAGGCAACCTTGGCTCACAGCTTTGTACTGAAGACTTTTGGCC 3197  
QY 2992 GAAGCCGATCCTTACATGATTAATTGATCAGAAATGTGTACAGAAACATACACTTGGCTT 3051  
DB 3138 CAGGCTCGATCTTCACTCTTCTATGATGAAGACACATTAACCATCTCTCAGTGGCTC 3257  
QY 3052 AAAGACATCAGAAATCCACGCTGAATTTTGGGATCCAGAAAGATGATCATAGTGA 3111  
DB 3258 TCCAGATGCAGAAAGACATGCGTGTTCAGAGAGCTGGGCTCACTGCTCAACATGCG 3317  
QY 3112 CTTCAGGTGCGCAATTAAGCTCAAGTACATTAACGCTATATTGTAACCTTCTCTG 3171  
DB 3318 ATTAAGGGAGGTGTAGAAATGAAGAGCCCTCTCGCTATGTACTATTCGCCCTTCTG 3377  
QY 3172 GGATATAGAAAGTATCAGCCTTAACATTAATGTGCAAGAGCTATCATTTTGGAGTCT 3231  
DB 3378 GAAATTCCTCTCTCCAGCTCACTAACCTTATGTTCGCAATGCCCTGTTCTGCCGAGTCA 3437  
QY 3232 GAATTCAGTAGAGAA 3247  
DB 3438 GCCTGAATGTAGCAA 3453

RESULT 12  
US-10-006-285-405  
Sequence 405, Application US/10006285  
Publication No. US20030165854A1  
GENERAL INFORMATION:  
APPLICANT: Mary Jane Cunningham  
APPLICANT: Matthew R. Kaeser  
TITLE OF INVENTION: MARKER GENES RESPONDING TO TREATMENT WITH TOXINS  
FILE REFERENCE: PA-0039 US  
CURRENT APPLICATION NUMBER: US/10/006,285  
NUMBER OF SEQ ID NOS: 514  
SOFTWARE: PERL Program  
SEQ ID NO 405

LENGTH: 4615  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: Incyte ID No. US20030165854A1 369213.39  
US-10-006-285-405

Query Match 3.0%; Score 126.8; DB 13; Length 4615;  
Best Local Similarity 51.6%; Pred. No. 2e-21;  
Matches 318; Conservative 0; Mismatches 292; Indels 6; Gaps 1;

QY 2638 TTGAGTTTCTCATTTCTCTCTAATACAGTACGTGCGAGTGAAGAGTTGACATCTGCA 2697  
DB 2838 TTGTCCTTGAAGCTCCCATCAATGTGTGCAAGATCTGCCAGAGCTTCTTTCAGTT 2897  
QY 2638 ATTGAGATGTTCTTGTGCTTCCATCAATGGCTTAGCCTCATTTGATTCGATCCTTAT 2757  
DB 2838 CTGGGTGACATATTATGTTCTCTCTATGCAAAATATACAAATCTCCTCAGATGCCATAT 2957  
QY 2758 GGCTGTGTAACAGAACATGATTAATTTGCTCCAAATATTACATTTTGGATTATCTG 2817  
DB 2958 GGCTGTGTAACAGAACATGATTAATTTGCTCCAAATATTACATTTTGGATTATCTG 3017  
QY 2818 ACTAAAGAAACAACTGACAGATTAATTGAAAGAAAGCTCTTTCATTATTAAGGCAA 2877  
DB 3018 AATGAACCCAGACAGTACGAGAGATCAAGGCCAAGCCGTTGGCTATCTCATCT 3077  
QY 2878 GGTTCACAGAGAACTTCTATACAGAGGAAGATGGCTCTTTCAGTCTTTGGG--- 2934  
DB 3078 GGTTCACAGAGAACTTCTATACAGAGGAAGATGGCTCTTTCAGTCTTTGGG--- 2934  
QY 2935 ---AATTATGACCCCTTCTGGAGCACTTGTGTGAGCTTTTGTATTAAGATTTTCTT 2991  
DB 3138 CGATATGGCAGGAACCAAGGCAACCTTGGCTCACAGCTTTGTACTGAAGACTTTTGGCC 3197  
QY 2992 GAAGCCGATCCTTACATGATTAATTGATCAGAAATGTGTACAGAAACATACACTTGGCTT 3051  
DB 3138 CAGGCTCGATCTTCACTCTTCTATGATGAAGACACATTAACCATCTCTCAGTGGCTC 3257  
QY 3052 AAAGACATCAGAAATCCACGCTGAATTTTGGGATCCAGAAAGATGATCATAGTGA 3111  
DB 3258 TCCAGATGCAGAAAGACATGCGTGTTCAGAGAGCTGGGCTCACTGCTCAACATGCG 3317  
QY 3112 CTTCAGGTGCGCAATTAAGCTCAAGTACATTAACGCTATATTGTAACCTTCTCTG 3171  
DB 3318 ATTAAGGGAGGTGTAGAAATGAAGAGCCCTCTCGCTATGTACTATTCGCCCTTCTG 3377  
QY 3172 GGATATAGAAAGTATCAGCCTTAACATTAATGTGCAAGAGCTATCATTTTGGAGTCT 3231  
DB 3378 GAAATTCCTCTCTCCAGCTCACTAACCTTATGTTCGCAATGCCCTGTTCTGCCGAGTCA 3437  
QY 3232 GAATTCAGTAGAGAA 3247  
DB 3438 GCCTGAATGTAGCAA 3453

RESULT 13  
US-09-981-151A-9  
Sequence 9, Application US/09981151A  
Publication No. US2003021256A1  
GENERAL INFORMATION:  
APPLICANT: Edinger, Shlomit R  
APPLICANT: Gerlach, Valerie  
APPLICANT: MacDougall, John R  
APPLICANT: Malyankar, Muriel M  
APPLICANT: Smithson, Glenda  
APPLICANT: Millet, Isabelle  
APPLICANT: Peyman, John A  
APPLICANT: Stoney, David J  
APPLICANT: Gunther, Erik  
APPLICANT: Ellerman, Karen  
APPLICANT: Shinkets, Richard A

APPLICANT: Padigaru, Muralidhara  
APPLICANT: Guo, Xiaojia  
APPLICANT: Patturajan, Meera  
APPLICANT: Taupier Jr, Raymond J  
APPLICANT: Burgess, Catherine E  
APPLICANT: Zernusen, Bryan D  
APPLICANT: Kekuda, Ramesh  
APPLICANT: Spytek, Kimberly A  
APPLICANT: Gangolli, Esha A  
APPLICANT: Fernandes, Elma R  
APPLICANT: Gorman, Linda  
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same  
FILE REFERENCE: 21402-168  
CURRENT APPLICATION NUMBER: US/09/981,151A  
PRIORITY FILING DATE: 2001-10-16  
PRIORITY APPLICATION NUMBER: 60/241,040  
PRIORITY FILING DATE: 2000-10-17  
PRIORITY APPLICATION NUMBER: 60/241,058  
PRIORITY FILING DATE: 2000-10-17  
PRIORITY APPLICATION NUMBER: 60/241,063  
PRIORITY FILING DATE: 2000-10-17  
PRIORITY APPLICATION NUMBER: 60/241,243  
PRIORITY FILING DATE: 2000-10-17  
PRIORITY APPLICATION NUMBER: 60/242,152  
PRIORITY FILING DATE: 2000-10-20  
PRIORITY APPLICATION NUMBER: 60/242,482  
PRIORITY FILING DATE: 2000-10-23  
PRIORITY APPLICATION NUMBER: 60/242,611  
PRIORITY FILING DATE: 2000-10-23  
PRIORITY APPLICATION NUMBER: 60/242,612  
PRIORITY FILING DATE: 2000-10-23  
PRIORITY APPLICATION NUMBER: 60/242,880  
PRIORITY FILING DATE: 2000-10-24  
PRIORITY APPLICATION NUMBER: 60/242,881  
PRIORITY FILING DATE: 2000-10-24  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 160  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 9  
LENGTH: 4488  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-981-151A-9

Query Match 2.9%; Score 125.2; DB 13; Length 4488;  
Best Local Similarity 45.6%; Pred. No. 5.2e-21;  
Matches 549; Conservative 0; Mismatches 628; Indels 27; Gaps 2;

QY 2049 TCACATGTCGGAAGCATTTTCCAGAGACTTGATTTGGCTAGACACCAACATGGTTA 2108  
DB 2187 TTCTCAGGTCCGCCAGTACTTCCAGAGACCTGGCTCGGATCTGTTCCATTTGGTAA 2246  
QY 2109 CAGATTTTACCAAGATTGGAAGTACGTACCTGATTTCTACTTTGGGGTGGCTAC 2168  
DB 2247 CTCGGGAAGGAGGGGTCCACAGTCCAGTTCGACCCCATCAACGAGTGAAGGCGAT 2306  
QY 2169 TGGTTTGTGATCTGAGAGACCTGGGCTTGAACATACTCCAGTGGAGCTTCA 2228  
DB 2307 GAGTTTCTGCACTTCCAGTCAAGAGGCTTGGGCTTTCAACCACT--GTTGACTTAC 2363  
QY 2229 AGCCTTCCACATTTTTCATTTTTCATCTTCCCTACTCTGTTATCAGAGGTGAAGA 2288  
DB 2364 TGCCTTCAAGCGCTTCTTGTGACCTGACCTCCCTTACTCAGTAGTCCGTTGGGAATC 2423  
QY 2289 ATTGCTTTGGAATACTATATTTCAATTTTGAAGATGCCACTGAGTTAAGTAT 2348  
DB 2424 CTTCCTGCTTACTGCAACATCTTCAATTAACCTAAAGATTGCAACAGGTTCAAGACTGA 2483  
QY 2349 CATGAGAAA-----GTGCAAAATTGATATTCTAATGAC 2384  
DB 2484 CTGCTTAATTCGATGAGTACCAAGTAAATCATGGGCAAGTTCTCAGACCTCCAGTTG 2543  
QY 2385 TTCAAGTAATTAATGCCAAGGCCACACAGACCACTTCTGTTCCAGTGAAGATGG 2444

DB 2544 TCCTGTGCTGATGACCCAAACCACACCTGGAAATCATCAGCTGTCAAATTTGGGTCA 2603  
QY 2445 GGCAACTGTTCTTTTCCATCAGGCCAACACATCTGGGAATTCCTTATCAGTCAAC 2504  
DB 2604 CATTAATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2663  
QY 2505 AGCTCTTTCACCCACTCTCTGATGCTGTACCCAGATGATTTTAAAGCTGAAG 2564  
DB 2664 GGGGTTTGTCCCAAAAGGCGCAAGTGAACGGCTCATCAGCTGCTGTCGCAACC 2723  
QY 2565 AATGAAAAATCATATTCATTCATCTTATTAAGTTGACTGACAAATAGGCTACAGAG 2624  
DB 2724 TGAAGGAGTCTGTGTGAAGAACACACACTTATGCTGTGCCAAAGAGAAAGGT 2783  
QY 2525 TACCCTGAAACTTGTGATTTTCTCATTTCTCCCTAATACAGTACCTGGAGTGAAGT 2684  
DB 2784 GGCATCTGAATCTGTCTCCCTGAGAGCTCCAGTGAACATTTGCTTACTGACCAAGC 2843  
QY 2685 TCAGATCACTGCAATTTGAGATGTTCTTGGTCTTCATCATATGGCTTACCTCATTTGAT 2744  
DB 2844 TTATGTTACGGTCTGTGGAGACATTAATGGGACAGCCCTGACAACTGGATGTTGCT 2903  
QY 2745 TCGATGCTTATGCTGTGTGTGAACAGAACATGATTAATTTGCTCAATATTTCAT 2804  
DB 2904 GCAATGCCAGTGGCTGTGGCGAGCAAAATGCTTGTGCTCCCATCATCTATGT 2963  
QY 2805 TTGGAATTAATCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2864  
DB 2964 CTTCAGTACCTGAGAGAGGAGGAGGCTGTGACGAGAGATCAGCTCTGGGAGATGGG 3023  
QY 2865 ATTATGAGCAAGTTTACAGAGAACTTCTATCAGAGGAAAGATGCTTTTCAG 2924  
DB 3024 TTTCCTGGAATTAAGGTATCCAGAGAGCTGATGTAACAAACAGCAATGGCTCATACAG 3083  
QY 2925 TGCTTTGGGAATTAATGACCTTCTGGGAGCACTTGTGTGACCTTTGTTTAAGATG 2984  
DB 3084 TGCTTTGGGAGAGGAGATGAAATGAAACATGCTCAGACGCTTGTTCAAATG 3143  
QY 2985 TTTCCTGGAAGCGATCTTACATAGATTAATTAATTAATTAATTAATTAATTAATTA 3044  
DB 3144 CTTTGGCCAAAGCTCAGAAATTCATCTTATTAATCCCAAGAACATCAGATGCTCTCA 3203  
QY 3045 TTGGCTTAAAGCATCAGAAATCCACAGGCTGAATTTTGGATCCAGAAAGATGATCA 3104  
DB 3204 GTGATGAGCAAGAACACAGCTCCCAAGTGGCTGTATGCCAAGTGGGAATCTCTTCA 3263  
QY 3105 TAGTGAAGCTTCAAGTGGCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3164  
DB 3264 CACAGCTATGAAGGAGTGTGATGATGATGAGTCTCTTGAAGTGTATGTCACAGCTGC 3323  
QY 3165 TCTCTGGGATATGAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 3224  
DB 3324 ATTGCTTGAAGTGGAAAGATGATGATGATGATGATGATGATGATGATGATGATG 3383  
QY 3225 GAGG 3228  
DB 3384 CAAG 3387

RESULT 14  
US-10-094-886-125  
Sequence 125, Application US/10094886  
Publication No. US20040002120A1  
GENERAL INFORMATION:  
APPLICANT: Kekuda, Ramesh  
APPLICANT: Tchernev, Velizar T.  
APPLICANT: Liu, Xiaohong  
APPLICANT: Spytek, Kimberly A.  
APPLICANT: Patturajan, Meera  
APPLICANT: Burgess, Catherine  
APPLICANT: Verneil, Corine A.  
APPLICANT: Li, Li



APPLICANT: Gorman, Linda  
APPLICANT: Malyankar, Uriel M.  
APPLICANT: Boldog, Ferenc  
APPLICANT: Guo, Xiaojia  
APPLICANT: Shenoy, Suresh  
APPLICANT: Padigaru, Muralidhara  
APPLICANT: Taupier, Raymond J., Jr.  
APPLICANT: Miller, Charles  
APPLICANT: Casman, Stacie  
APPLICANT: Pena, Carol  
APPLICANT: Gangolli, Esha  
APPLICANT: Gusev, Vladimir  
APPLICANT: Smithson, Glenda  
APPLICANT: Zeinusen, Bryan  
APPLICANT: Gerlach, Valerie  
APPLICANT: Pochart, Pascal  
APPLICANT: Fernandes, Elma  
APPLICANT: Shinkets, Richard  
APPLICANT: Rastelli, Luca  
APPLICANT: Spaderna, Steven  
APPLICANT: Larochele, William  
APPLICANT: Zhong, Mei  
TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD  
FILE REFERENCE: 21402-290 B  
CURRENT APPLICATION NUMBER: US/10/094,886  
PRIOR FILING DATE: 2002-03-07  
PRIOR APPLICATION NUMBER: 60/274,322  
PRIOR FILING DATE: 2001-03-08  
PRIOR APPLICATION NUMBER: 60/313,182  
PRIOR FILING DATE: 2001-08-17  
PRIOR APPLICATION NUMBER: 60/288,052  
PRIOR FILING DATE: 2001-05-02  
PRIOR APPLICATION NUMBER: 60/318,510  
PRIOR FILING DATE: 2001-09-10  
PRIOR APPLICATION NUMBER: 60/274,281  
PRIOR FILING DATE: 2001-03-08  
PRIOR APPLICATION NUMBER: 60/314,018  
PRIOR FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: 60/274,194  
PRIOR FILING DATE: 2001-03-08  
PRIOR APPLICATION NUMBER: 60/274,849  
PRIOR FILING DATE: 2001-03-09  
PRIOR APPLICATION NUMBER: 60/296,693  
PRIOR FILING DATE: 2001-06-07  
PRIOR APPLICATION NUMBER: 60/313,626  
PRIOR FILING DATE: 2001-08-21  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 298  
SOFTWARE: PatentIn 2.1  
SEQ ID NO 125  
LENGTH: 4501  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(4309)  
US-10-094-886-125  
Query Match 2.8%; Score 119.8; DB 12; Length 4501;  
Best Local Similarity 45.7%; Pred. No. 1.3e-19;  
Matches 596; Conservative 0; Mismatches 682; Indels 27; Gaps 4;  
QY 2040 GGATGACAGTCCAGATGTCGAAGCATTTTCCAGAGATTTGGCTAGACACCA 2099  
DB 2088 GGCTATATTCACAGTAAGAACAACTTCCAGAGATGATGAGACCTGTCAG 2147  
QY 2100 CATGGTTACAGATTTACAGAAATTTGAAGTAAGTGAATCTATCATCTTCTG 2159  
DB 2148 TGTGATTCCTCAGGCTCGCAATCTTCTGTTCCATTCCTGATACATTAACCAAG 2207  
QY 2160 GGTGCTACTGTTTGTGATCTCTGAGAGCTGGGCTTTGAGCTAACAATCTCACT 2219  
DB 2208 GGAGGCAAGTGGCTTTGTGTGATGTGACGTTGATTTGGCATTTCTCTAACAACAC 2267

QY 2220 GGAGCTCCAGGCTTCCAAACCATTTTTCATTTTGTGATCTTCCCTACTCTGTATCAG 2279  
DB 2268 T---CTAGAAGCTCCCAACCTTCTTATTAAGATGCGTCACCTTTTGGTGTTC 2324  
QY 2280 AGGTGAAGATTTGCTTTGAAAATACATATATTTCAATTTTGAAGATGACATGAGT 2339  
DB 2325 AAATGAACAATTTGATTTGATTTGATGATGATGCTTACGAGATATCATGTTGAGAT 2384  
QY 2340 TAAAGTATCATTTGAGAAAAGTGAACAATTTGATTTCTAATGACTTCAAGTGAATAA 2399  
DB 2385 TTCTGTCAAGTGGAGAGTCTCAGAAATTAAGCAAAATATTCATCTTGAATCA 2444  
QY 2400 TGCCAGGCCCACGACAGACCTTCTGATTTCCAGTGAAGATGGGGCACTGTCTTT 2459  
DB 2445 TGCGATGAGTTATTTCAAGCTGAGGAGGAGAAAACAGCTGTGATATATACCTAA 2504  
QY 2460 TCCCATCAGGCCAACACATCTGGGAAATTCCT-ATCACAGTCAAGCTTTTACCCA 2518  
DB 2505 GAATTTGGGTAAAGATATATCATGTTGTTGCTGATGCCAAACAGAGTGTGCC 2564  
QY 2519 CTGCTTGATGCTGTCAACAGATGATTTTATGTAAGCTGAAGGATAGAAAATCAT 2578  
DB 2565 AAATGAAGAAATGAGCAGCAAAACCTAAGTGAAGACATGTGTCCAAAGCTTCTT 2624  
QY 2579 ATTCACAATCCATCTTATTTAGACTTGAACAATAGGCTCAGAGTACCCTGAACCTT 2638  
DB 2625 AGTAAAGCTGAAGTATTTGAAAAGAAAAGAACCCAGAGTTCTTATCTGTACGAAGG 2684  
QY 2639 TG-----AGTTCTCATTTCTCTATATACAGTACGTGACGTGA 2678  
DB 2685 TGCCAAAGCTTCCAGCAGGAGGTTTGAACTTGCACAAACGATATGATGAAGGGTCAAG 2744  
QY 2679 AAGATTTAGATCATCTGAATTTGAGATGTTCTTGTCTTCAATAGGCTTACGCTC 2738  
DB 2745 CAGAGCTTTTCACTGTGTGGGGGATATCTAGAGATTTGCTTCCAGAAATCTG3--T 2801  
QY 2739 ATTGAATGGAATGCTTATGCTGTGTGTAACAGAAATGATTAATTTTGTCCAAATAT 2798  
DB 2802 TGTTCCTCAAAATGCTTATGGAAGTGAAGAGCAATGCTGCCCTATGATGATGATG 2861  
QY 2799 TTACATTTTGAATTTATCTGACTAATAAAGAAACATGACAGATTAATTTGAAAGAAAG 2858  
DB 2862 TTATGTTTGGACATATCTGAATATCTAGAGAACTGACAGAGAAATTTCAATCTAAGG 2921  
QY 2859 TCTTTCAATTTATGAGGCAAGTTTACAGAGAACTTCTATCAAGAGAAAGATGCTC 2918  
DB 2922 TTTCTTCTCTATCTAATGATGTTATCAAGGCAATATCTTCAAAAACCTGATGCTTC 2981  
QY 2919 TTTGAGTCTTTGGGAATTTATGACCTTGGGGAGCACTTGTGTGACGTTTGTGTTT 2978  
DB 2982 CTATATGTTGTTTGGAGCAGAGTCAAGAAAGAAAGATGCTGATGCTCTTATCTT 3041  
QY 2979 AAGATGTTCTTGAAGCCGATCTTATCATGATATGATGATGATGATGATGATGATGATG 3038  
DB 3042 TAAAGCATTTGAGAGAAATGAATAATATGATATGATGATGATGATGATGATGATGATG 3101  
QY 3039 ATACATTTGCTTAAAGACATCAGAAATCCAAAGCTGAATTTTGGATCCAGAAAGT 3098  
DB 3102 CTATATGCTGCTTCAACCAACAGAAACAGGCGCTTTTAAAGATGATGAGCAGCT 3161  
QY 3099 GATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3158  
DB 3162 TTTCAACACAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3221  
QY 3159 AACTTCTCTCTGAGATATGAAGATATCAGCTTAACATGATGATGATGATGATGATGATG 3218  
DB 3222 TGGGATGTTCTTTGAAGCTGGGCTCAATTTTCACTTTCTGTGTCTGAGAAAGCAGCTTT 3281  
QY 3219 TTTTGTGAGTCTGAATTCAGTGAAGATTTCAAGCAATTTATCTAGGCTTATTAAC 3278  
DB 3282 TTGCTTGAAGGAGGATTTGACAGTGTGTGATGATGATGATGATGATGATGATGATGATG 3341





PUBMED  
REFERENCE  
AUTHORS  
11042159  
3  
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,  
Kono,H., Akiyama,J., Nishi,K., Kitsuana,T., Tashiro,H., Itoh,M.,  
Sunb,N., Iehi,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,  
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,  
Fujikawa,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watabiki,M.,  
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J.,  
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
11076861  
4  
Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,  
Aizawa,T., Hara,A., Fukunishi,Y., Kono,H., Adachi,T., Fukuda,S.,  
Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamataka,I.,  
Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,  
Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,  
Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Koichiwa,H.,  
Kuehl,P., Lewis,S., Matsuo,Y., Nikaide,I., Pevsley,G., Quekembush,J.,  
Schriml,L.M., Staudt,F., Suzuki,R., Tomita,M.,  
Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,  
Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,  
Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,  
Fleischer,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,  
Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,  
Marchionni,L., Mashima,J., Mazzarelli,J., Nombiers,P., Nordone,P.,  
Rising,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,  
Sato,K., Schonbach,C., Seiya,T., Shibata,Y., Storch,K.F., Suzuki,H.,  
Toyooka,K., Wang,K.H., Weitz,C., Whitaker,C., Wilming,L.,  
Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawai,H., Kohlski,S.  
and Hayashizaki,Y.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409 (6821), 685-690 (2001)

JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
11217851  
5  
The FANTOM Consortium and the RIKEN Genome Exploration Research  
Group Phase I & II Team.  
Analyses of the mouse transcriptome based on functional annotation  
of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

JOURNAL  
REFERENCE  
AUTHORS  
6 (bases 1 to 4485)  
Adachi,J., Aizawa,K., Akimura,T., Arawaka,T., Bono,H., Carninci,P.,  
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,  
Hayashida,K., Hayatsu,N., Hiramoto,K., Hirooka,T., Hirozane,T.,  
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,  
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Kono,H., Kouda,M.,  
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,  
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,  
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,  
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shitaki,T.,  
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaki-Akahira,S.,  
Takeda,Y., Tanaka,T., Tomaru,A., Toyu,T., Yasunishi,A.,  
Muramatsu,M. and Hayashizaki,Y.  
Direct Submission  
Submitted (16-JUN-2001) Yoshihide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp,  
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,  
Fax: 81-45-503-9216)  
cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in RIKEN  
Genomic Sciences Center and Genome Science Laboratory in RIKEN  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.  
Please visit our web site for further details.  
URL: http://genome.gsc.riken.go.jp/  
URL: http://fantom.gsc.riken.go.jp/.

FEATURES  
source  
Location/Qualifiers  
1..4485  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM\_DB:4831440K17"  
/db\_xref="taxon:10090"  
/clone="4831440K17"  
/issue\_type="head"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="0 day neonate"  
misc\_feature  
112..4438  
/note="putative  
weakly similar to alpha-2-macroglobulin (fragments)  
[Limulus polyphemus] (PIR|A36260, evidence: FASTY,  
51.8%ID, 76.7%length, match=168)"

BASE COUNT  
1277 a 970 c 1010 g 1228 t

ORIGIN  
Query Match 63.2%; Score 2708.8; DB 11; Length 4485;  
Best Local Similarity 77.9%; Pred. No. 0;  
Matches 3393; Conservative 0; Mismatches 872; Indels 91; Gaps 7;  
1 ATGCGAGGGCCACCGCTCTGACCGCCGACCTCTCTGCTGTCACCGCCGCTG  
112 ATGCGAGGTCGAGGCTCTGACCGCCGACCTCTCTGCTGTCACCGCCGCTG  
61 GCGGTGCTCCCGGCTCGGTTTCTGGTGACAGCCCGAGGATCATCAGCCCGAGCA  
172 GCC--GCCCGAGGTCCCGCTTCTGTGTACAGCCCGAGGATCATCAGCCCGAGCA  
121 AATGTGACTATTTGGGTGAGCTTCTGGAACACTGCTTCTGACAGGTGATGGAAGCG  
229 AATGTGACTATTTGGGTGAGCTTCTGGAACACTGCTTCTGACAGGTGATGGAAGCG  
181 GAGCTGCTCAAGACAGCATTAACCTCAGCTCTCTGCTGTCGAGCAGAGAGTCTT  
289 CAGGTGTTCAAGATAGCTTCAACCAATCGATTCATCTGGAAGCAGAGAGTCTT  
241 GAAAGAGGCTTTTAAACACTTACCTTCCATCACTACCTCGAAGCAGTGAAGT  
349 CACGAGGCGATTTCAAGACTCTGCTTCTGCGGCACTCTGACACATGCGATGAG  
301 ATTATGAGCTACGTTGATCCGAGCAGTACCCAGATGAGATTTATTTCTTAATGATCC  
409 ATTATGAGCTACGTTGATCCGAGCAGTACCCAGATGAGATTTATTTCTTAATGATCC  
361 CGCTTATCATTTGAGACCAAGAAATATCTGCTTCAATTAACAGCAAGCCTTATAC  
469 CGTTTACATTTGAGACCAAGAAATATCTGCTTCAATTAACAGCAAGCCTTATAC  
421 AAGCGAAGCAAGAAAGTATTCGATGTTACACTCTTCCAGATTTAAGCCTTAC  
529 AAGCGAAGCAAGAAAGTATTCGATGTTACACTCTTCCAGATTTAAGCCTTAT  
481 AAAACCTCTTAAACATTTCTATTAAGACCCCAATCAATTTGATTCGAACAGTGTG  
589 AGGACATCCGTGACATTTCTATTAAGATCCGAAGTCTAATGATCCGAAGTGTG  
541 TCACCAAGAAAGTATTTGAGAGTCAATTTCCAAATCTTCACTATCTTCCATCAATA  
649 TCTCAGAAAGTATTTGAGAGTCAATTTCCAAATCTTCACTATCTTCAATCAATA  
601 CTGTGACCTGCTATTTCAAGTTCAGAGCAAGACCAAGCATATTAATCAATTTTCA  
709 TTGGTACCTGCTATTTCAAGTTCAGAGCAAGACCAAGCATATTAATCAATTTTCA  
661 GTTCAAGATATGATTTCAAGATTTCAAGATTTCAAGATTTCAAGATTTCAAGATTT  
769 GTCTAGAAATATGTTTCAAGATTTCAAGATTTCAAGATTTCAAGATTTCAAGATTT  
721 ATGAATTTAAGATTTAAGATTTCAAGATTTCAAGATTTCAAGATTTCAAGATTT

Db 829 TTGAAGCTTAAACGTTAAATGCTCCGTCATAGCAAAAGTATACATACGGGAAGCCAGTGG 888  
 Qy 781 AAAGAGACGTACCGTTACATTTTACCTTTACCTTTTGGGGAAGAAAGAAATATAT 840  
 Db 889 AAGGGAAGTCTGTCACTTACATTTTACCTTTGTCTTTTGGGGAAGAAAGAAATATAC 948  
 Qy 841 AAAAAACATTTAAGTAAATGATGTGCAAACTTCTCTTTTAATGATGAAGATGAA 900  
 Db 949 AAAAAAGTTTGAAGTAAATGATGTGCAAACTTCTCTTTTAATGATGAAGTAA 1008  
 Qy 901 AATGTAT-----GATTTCTTCAATGACTTTCTGATATCCGTAT 942  
 Db 1009 AAGGTATGATTTGAAACCGCTACGATGCTCGAAGGAGTTATGAAATGTGGAC 1068  
 Qy 943 CTATCTTCCCTGACCGACGTAGAAATTTTACCAAGTACAGAAATCAGTTACAGTAT 1002  
 Db 1069 CGTCTTCTTGGGCGGCGAGAGTATTAAGCTACGTACCGAAATCATCACAGTATC 1128  
 Qy 1003 TCAAGAAATGTAGCACTAATGTGTTCTTCAAGCAACATGATTAATCATGATTTT 1062  
 Db 1129 TCAAGATGCGAAGCAACATGTGTTTCAAGCAACATGATTAATCATGATTTT 1188  
 Qy 1063 GATTAATCTACTGTCTTGAAGCAGTCTCAACTTCAAGCCACTGTGAAGTACTGT 1122  
 Db 1189 GATTAATGCACTGTCTTGAAGCAGTCTCTCAACTTCAAGCCACTGTGAAGTACTGT 1248  
 Qy 1123 GCTGATGCGAACCACCTGACTCTTGAAGAAAGAAATATGATGCTAATACAGTACA 1182  
 Db 1249 TCTGATGCGAATCACTAATCTCTGAGAAATGAAATGATTTGTCACAGTATGACC 1308  
 Qy 1183 CAGAGAACTATATCTGAGTACGTGAGCGGATCTAACAGTGAATCAGAAAATGAAAGCT 1242  
 Db 1309 CAGGGAATAATTAACCACTGAGAGTACGA-----GAGATCAGAAAATGATTAAT 1358  
 Qy 1243 GTTCAAGAAATTAATTAATCTGTCCCGCAAGTGAAGTCTTTAAGTTGAATTTCCCAATC 1302  
 Db 1359 ATCCAGAACCGTAATTAATTAATCTGTCCCGCAAGTGAAGTCTTTAAGTTGAATTTCCCAATC 1418  
 Qy 1303 CTGAGAGATTTCCAGTACGTACAGTTGAAGGCTTATTTCTGTTAGTAAAGTATGACATG 1362  
 Db 1419 ATGTCTATTTCCGGCGAGCTTACAGTGAAGGCTTATTTCTGTTAGTAAAGTATGACATG 1478  
 Qy 1363 GCAGTTCAATAGTCTGTTTAAGTCTCTAGTAAAGACATACATCCAACTAATAAACAAGAT 1422  
 Db 1479 ACAGTCCACAGTATGTTACCTCTCTAGTAAAGATGATCATCAGTTTAAACAGAGAT 1538  
 Qy 1423 GAAATATTAAGTGGAGTCCCTTTGAGTTGGTGTGTTAGTGGCAACAAGATGAAG 1482  
 Db 1539 GAATATATTAAGTGGAGTCCCTTTGAGTTGGTGTGTTAGTGGCAACAAGATGAAG 1598  
 Qy 1483 GATTAAGCTATATGATGATATCCAGGGGACAGTTGGGCTGTAGGAAACAATAATCA 1542  
 Db 1599 GACTTAAGCTATATGATGATATCCAAAGGACAGTTGGGCTGTAGGCAACAATAATCA 1658  
 Qy 1543 ACAATGTTCTCTTTAAACACAGAAATTTCTTGAAGTCAAAAGCCTGTGTAATTTGTAAT 1602  
 Db 1659 ABAATCTTCTCTTTAAACACAGAGGCTTCTGGGCTCCAAAGCCTGTGTAATTTGTAAT 1718  
 Qy 1603 TATATTAAGATGATGGGAAATTAATAGTATGTTTAAATAATTTCTGTTCACTGTT 1662  
 Db 1719 TATATTTGACAGAGATGGGAAATTAATAGTATGTTTAAATAATTTCTGTTCACTGTT 1778  
 Qy 1663 TTTAAATAATTAAGTAAAGCTATATGATGATGAAAGTGAAGCCATCTGAGAAAGTC 1722  
 Db 1779 TTTGAAATAATTAAGTAAAGCTATGTTGAGATTAACCTACAGTCAAGCCATCTGATTAAGTC 1838  
 Qy 1723 TCTTTAGATCTCTGTGACACAGCTGATCCATAGTTGGATGTTAGCTGTTGCAAA 1782  
 Db 1839 TCTCTCAGGATCTCTGCGACACAGTGTGCTCCGTGGGGATGTTAGCTGTTGCAAA 1898  
 Qy 1783 AGTGTGATCTGTGATGATGCTCTTAATGATTAATTAATGAAAGTGGTCCATGATGG 1842  
 Db 1899 AGTGTAACTGATGAAATCTCCAAACAGTATTAATGAAAGTGGTCCATGATGATGG 1958

Qy 1843 GAACCTTAAACACAGATATATTTAGCATGTTTCATGAATTTCTTTGACAGTCTTTCAG 1902  
 Db 1959 GAACCTTAAACACAGATATATTTAGCATGTTTCATGAATTTCTTTGACAGTCTTTCAG 2018  
 Qy 1903 GAATGTGACCTGTGGGTATTTGACAGATGCAAACTCACAGAGATTAATTTAGTGTGTT 1962  
 Db 2019 GAATGTGTCTGTGGGTATTTGACAGATGCAAACTTCATTAAGATGATGATGATGTC 2078  
 Qy 1963 TATGACATGACGAATATGCTGAGAGTGTGAGGAAATGAAAGACATATGTATAT 2022  
 Db 2079 TACGATCTGAAAGATATTTGAAAGCTTTGCAAGGAAATGAGCAACCTGATAT 2138  
 Qy 2023 ATTCAATGCTTTCTTTGGGTAGACATCCATGTCGAAAGATTTTCCAGACATTTG 2082  
 Db 2139 TTTGAAGATGCTTCTTCAAGTATGATGATGATGATGATGATGATGATGATGATGAT 2198  
 Qy 2083 ATTTGGCTAGACACCAATGAGTTTACAGATTTTACAAAGATTTGAAGTATGATGAT 2142  
 Db 2199 ATTTGGCTAGACACCAATGAGTTTACAGATTTTACAAAGATTTGAAGTATGATGAT 2258  
 Qy 2143 GATTTATCACTTCTTGGGTGCTATCTGTTTGTGATCTCTGAGACCTGGCTTGGGA 2202  
 Db 2259 GACTCATCATCTTCTTGGGTGCTATCTGTTTGTGATCTCTGAGACCTGGCTTGGGA 2318  
 Qy 2203 CTAAACATCACTCCAGTGGAGCTCCAGCCCTTCCAACTTTTCAATTTTGTGATCT 2262  
 Db 2319 TTAACACCGTTCCAGCAGAGCTGACAGCTTCCAACTTTTCTTTCTGATATCT 2378  
 Qy 2263 CCTACTCTGTTATCAGAGGTGAAGAAATTTCTTGAATTAATTAATTAATTTG 2322  
 Db 2379 CCGTACTCTGTTATCAGAGGTGAAGAAATTTCTTGAATTAATTAATTAATTTG 2438  
 Qy 2323 AAAATGCACTGAGGTAAAGTATCAATGAAAGTGAAGAAATTTGATTAATTTGATG 2382  
 Db 2439 AAAATGCACTGAGGTAAAGTATCAATGAAAGTGAAGAAATTTGATTAATTTGATG 2498  
 Qy 2383 AACTCAATGAAATTAATGCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGAT 2442  
 Db 2499 AACTCAATGCAACCAATGCAACCAATGCAACCAATGCAACCAATGCAACCAATGCAAC 2558  
 Qy 2443 GGGGCACTGTTCTTTTCCCATGAGGCAACATCTGGGAGAAATTTCTATCACTG 2502  
 Db 2559 GGGGCACTGTTCTTTTCCCATGAGGCAACATCTGGGAGAAATTTCTATCACTG 2618  
 Qy 2503 ACAGCTCTTTCAACCACTGCTTCTGATGCTGTCAACCAAGATTTTATGTAAGCTGAA 2562  
 Db 2619 ACGGCTGCTCAACCACTGCTTCTGATGCTGTCAACCAAGATTTTATGTAAGCTGAA 2678  
 Qy 2563 GGAATGAAATAATCAATTTCAATCAATCTTATTTAGCTTGAAGTGAAGTGAAGTGAAG 2622  
 Db 2679 GGAATGAAATAATCAATTTCAATCAATCTTATTTAGCTTGAAGTGAAGTGAAGTGAAG 2738  
 Qy 2623 AGTACCTGAAATAATTTGATTTCTATTTCTCTATATACAGTACAGTGAAGTGAAGTGAAG 2682  
 Db 2739 AGTAAAGCAATCATGATGATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2798  
 Qy 2683 GTTCAATCACTGCAATTTGAGATGTTTGTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 2742  
 Db 2799 GTTCAATCACTGCAATTTGAGATGTTTGTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 2858  
 Qy 2743 ATTTGGATGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2802  
 Db 2859 ATTTGGATGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2918  
 Qy 2803 ATTTGGATGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2862  
 Db 2919 ATTTGGATGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2978  
 Qy 2863 TCAATTTAGGAGCAAGTTACAGAGAACTTCTATCAAGAGGAAAGTGGCTCTTTC 2922  
 Db 2979 TCAATTTAGGAGCAAGTTACAGAGAACTTCTATCAAGAGGAAAGTGGCTCTTTC 3038

OY	2293	AGTGGTTTGGGAATTATGACCCCTCTG36AGCACTGTGGTGCAGCTTTGTTTAA	2982
Db	3039	AGTGGTTTGGGCAATTGACTCTTCTGGAGACCTTGGCTGTCACTTTGTTTAA	3098
OY	2983	TGTTTCTCTTGAAGCCGATCTTACATAGATATTGATTCAGAAATGTGTTACAGACATAC	3042
Db	3099	TGCTTTCTGGAAAGCTGATTACTATATATGATTTGATTCAGAAATGTGTTACAGAAATAT	3158
OY	3043	ACTTGGCTTAAAGACATTCAGAAATCCACGGTGAATTTTGGGATTCAGAAAGTGA	3102
Db	3159	ACTTGGCTTAAATGCACATPAGAAATTCATGTGTGAATTTTGGAGCCAGAAAGTGA	3218
OY	3103	CATATGAGCTTCAAGGTGGCAATTAAGTCCAGTAACACTTACGCTTAATATGTACT	3162
Db	3219	CACATGTAACTTCAAGGTGGCACCAAAGCCAGTAACTTACCGCTTAATATGTGACT	3278
OY	3163	TCTCTCTGGGATATATGAAGAAGTATCAGCTTAACTTGAATGTGCAAGACTTATCCATTT	3222
Db	3279	TCTGTCTGGGATATCAAAAAGTATCAGCTTAATATCGATGTATCAAGACTCAAGTT	3338
OY	3223	TTGAGCTGAATTGATGAGAAATTTCAACAATTAATCTAGCCCTTAATACTTAT	3282
Db	3339	TTGGAATTTGAATTGAGAGAGAAATTTGGCACTATTAACCTTAGCAATTAATCTTAT	3398
OY	3283	GCATTGTCAATCAGTGGGGAAGTCTTAAAGCAAGAAAGCTTTGAATATCTGACTTTGGA	3342
Db	3399	GCCCTGTCCACAGTGGGAGCCCTTAACAGAGGAGCTTTGAACTTCTATGACGA	3458
OY	3343	GCAGAAACAAGAAAGTGGCATGTCGAATTTCTGGGTCATCAGAGTCCAACTTCTGACTCC	3402
Db	3459	TCAGAAAAGAAAGAGACACACAGTTCTGTGTGTATCAGAGACCTGGCTCTGTGGTCC	3518
OY	3403	TGGACGCAAGCTCCCTGGATATTGAAGTTGCAAGCTTATGCACTGTCTCACTTCTTA	3462
Db	3519	TGGCAGCGCGCTCCGTGGACATCGAAATCGCTGTATGCACTGTGGCGCACACTGTG	3578
OY	3463	CAATTTGACACTTCTGAGGAAATCCCATATATAGAGTGGCTTAAGACGCAAGAAATATGC	3522
Db	3579	CA---CAATGTGTCTGAGGGAATCCCGGTTATGAGTGGCTATCCACCAAGAAACAGC	3639
OY	3523	TTGGGTGGTTTTCATCTACTACGAGATPCCACTGTGGCTTTAAAGGCTCTGTCTGAATTT	3582
Db	3636	CTGGAGGTTTGTATCCACTCAGATATCTGTGTGGCCCTGAGAGCAATTAATCTGAATTT	3699
OY	3583	GCAGCCCTAATGAATACAGAAAGCAAAATATCCAATGACCGTACCGGCGCTTAGCTCA	3642
Db	3696	TCAGCCCTTGTGCAAGAGAAATATCAGATATCCAATGACCGTACAGGGCCGGCATCC	3755
OY	3643	CCAACTCT-----	3651
Db	3756	CCCAAGATCCATPACCTTCCGAATTCCTCCAGAACTCTTCTCTCTTCAACAGAAAG	3815
OY	3652	CTTGTGTGTGTAACAGCAATGSCAGCTTAATATTTCCGAATATGTTTGAATTTGATTT	3711
Db	3816	CTTATAGGCTAGATCCCATACCGTATATGTTTCTGGCAATGGCTCGGAAATTTGTA	3875
OY	3712	TGTCACTCAATGTGTATATATATATGAAAGCTTCTGGCTCTTCTAGAAAGCAAGATCT	3771
Db	3876	TGCAAGCTTAATGTGATCAATPACCTGAAAGTTAGGTTCTTCTAAAGACGAGATCT	3935
OY	3772	ATCCAAATCAAGAGCCTTGTATTTAGATGTTCTGTAAAGAAATTAAGATGATCTC	3831
Db	3936	ACCGAAATCAAGAGATTTTGAATTTAACGTCATTTGA---ATTAAGAGACGACATTT	3992
OY	3832	AATCATGTGATTTGAATGTGTATCAAGCTTTTGGGCCCCGGATAGAGATGGCT	3891
Db	3993	AGTCACTGAATCTGAATGTGTGCAAGTCACTTGGGTTTCAAGAGGACAGCAATGCTC	4052
OY	3892	CTTATGAGATTTAACTTATTAAGTGGCTTTATATGTCCTTCAAGAAACAATTTCTGAGC	3951
Db	4053	CTTATGAGAGTGAACCTTCTCAGTGGCTTTATGTATCATTTCAAGATTCATTTCTCTGAGT	4112
OY	3952	GAGACAGTGAAGAAAGTGAATATGATCATGAGAAACTCAACTTATTTAGATTTCTGTA	4011

Db	4113	GAGACCTGAAAGAAAGGGAATATGATTAACGGGAAATCAACTTTATTTAGATCTGTGA	41172
Qy	4012	AATGAAACCCAGTTTGTGTATATTTCTGCTGTGAGAACTTTAAAGTTTCAATATAC	40711
Db	4173	AATGAATCCAGTTTGTGTATATTTCCACTGTGAGAACATCAAAAGTTTCAAAATTT	42322
Qy	4072	CAAGATGCTCAGTGTCCATAGTGGATTACTATAGAGCAAGGAGACAGGGCGTAGAAGT	41311
Db	4233	CGAATATGTTCAAGTATCTGTAAATGACTTACTATAGAACCAAGAGACAGGCACTGTCCAAAG	42922
Qy	4132	TACAACTCTGAAGTGAAGCTGTCTCTCTGTACCTTTCAGTCAGTATGTCCAGGGCTGCGGT	41911
Db	4293	TACAAACCCAGGTGAAGCTGTCTCATGTATTACT--CAGTCCAGACACCAACTGCAAG	43492
Qy	4192	CTTTGTAGAGATGAGACTTCAAGGCTCCCATCATCACTTCCTCAGTCATTTTATTTTCTGT	42511
Db	4350	TCACACACGAGCGAGCGACGAGACTCCCTTCGAGCTTCCTTCGAGCCTTCTGTCTTTTGC	44092
Qy	4252	TTCAAGCTCTGTATCTTATGGAACCTTTGGCTGTGA	42872
Db	4410	TCGCTCTCTGTACTTTGTGCAACATTTATATGA	44415

LOCUS	4725 bp	RNA	linear	HTC 05-DEC-2002
AK036799				
DEFINITION	Mus musculus adult female vagina cDNA, RIKEN full-length enriched library, clone:993012E15 product:weakly similar to alpha-2-macroglobulin (fragments) [litmus polypheumus], full insert sequence.			
ACCESSION	AK036799	GI:26085434		
KEYWORDS	HTC; CAP trapper.			
ORGANISM	Mus musculus	(house mouse)		
REFERENCE	Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
AUTHORS	1 Carninci, P. and Hayashizaki, Y.			
TITLE	High-efficiency full-length cDNA cloning			
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)			
MEDLINE	99279253			
PUBMED	10349636			
REFERENCE	2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.			
AUTHORS	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes			
TITLE	Genome Res. 10 (10), 1617-1630 (2000)			
JOURNAL	20499374			
MEDLINE	11042159			
PUBMED	3			
REFERENCE	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Komno, H., Akiyama, J., Nishi, K., Kitsuaki, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishire, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujikawa, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Wataniki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.			
AUTHORS	RIKEN integrated sequence analysis (RISA) system-384-format			
TITLE	sequencing pipeline with 384 multicapillary sequencer			
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)			
MEDLINE	20530913			
PUBMED	11076861			
REFERENCE	4			
AUTHORS	Kawai, J., Shingawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arawaka, T., Hara, A., Fukunishi, Y., Komno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamataka, I., Saito, T., Okazaki, Y., Gotojohri, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleisheimann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kiehl, P., Lewis, S., Matsuo, Y., Nakido, I., Pessio, G.,			

Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M., Wagner, T., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldacci, L., Barish, G., Blake, J., Hoffell, J.D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hochmann, L., Hume, D.A., Kamuya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzerelli, J., Mombascher, P., Nordone, P., Ring, B., Rindwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyokura, K., Wang, K.H., Weitz, C., Whitaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohetsuki, S. and Hayashizaki, Y.

Functional annotation of a full-length mouse cDNA collection  
*Nature* 409 (6821), 685-690 (2001)

JOURNAL MEDLINE  
PUBMED  
11055660  
11217851

5

The FANTOM Consortium and the RIKEN Genome Exploration Research

REFERENCE  
AUTHORS

**TITLE**  
Analysis of the mouse transcriptome based on functional annotation  
of 60,770 full-length cDNAs  
**JOURNAL**  
Nature 420, 563-573 (2002)  
**REFERENCE**  
(bases 1 to 4725)  
**AUTHORS**  
Aachi, U., Aizawa, K., Akimura, T., Arahawa, T., Bono, H., Carninci, P.

**AUTHORS**  
Adachi, J., Aizawa, K., Akimura, T., Aikawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Hori, P., Imocani, K., Ishii, Y., Itoh, N., Kagawa, I., Kaubuka, T., Katoh, H., Kawai, J., Kojima, Y., Komoto, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Nunazaki, R., Ohno, M., Ohshio, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sugabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toy, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.

**TITLE** Direct Submission  
**JOURNAL** Submitted (16-JUL-2001) Yoshinide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.go.jp, URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

**COMMENT** cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
Please visit our web site for further details.  
URL: <http://genome.gsc.riken.go.jp/>  
URL: <http://fantom.gsc.riken.go.jp/>.

FEATURES	Location/Qualifiers
source	1. .4725

misc\_feature

/note="putative  
weakly similar to alpha-2-macroglobulin (fragments)  
[Limulus polyphemus] (PIR|A36560, evidence: FASTY,  
1988)

polyA\_signal  
polyA\_site

	/note="putative
BASE COUNT	1334 a 1016 c 1066

BASE COUNT	1334 a	1016 c	1066 g	1309 t
------------	--------	--------	--------	--------

## ORIGIN

Query Match	62.3%;	Score 2668.8;	DB 11;	Length 4725;
Best Local Similarity	77.7%;	Pred. No. 0;		
Matches 3389;	Conservative	0;	Mismatches 877;	Indels 94;
				Gaps 10

OY	1	ATGCAAGGGCCACACGGCTCTTACACGGCCGCCACCTCTCTGAGTGTGACCGCCGCTG	60
Db	97	ATGCGCASTCGAAGGCTCTCTAGCGCGGCCACACTTCTCTGCTTGTGGCGGTCCGCGCTG	156
OY	61	GCCGTGCTCCCGG3CCTCGGTTTCTGTGTGACAGCCCCAGGGATCATGAGCCCGGAGGA	120
Db	157	GCC---GCCCAGGGTCCCGCTTCTGTGTGACAGGCCAGGGATCATGAGCCCGAGGA	213
OY	121	AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCAACAGTGACTGTGAAGGCG	180
Db	214	AATGTGACTATTGGGGTGGAGTCTCTGTGAAAATAGCCCCCACAAGCTCTGTGTAAAGGT	273
OY	181	GAGCTGCTCAAGACAGCATCAAACTCACTGTCTGTGCTGTGAAGCAGAAGAGTCTTT	240
Db	274	CAGGTGTTCAAGATAGCTTCCAACTGAGATTCATCTCTGGAAGCAGAAAGAGTCTTT	333
OY	241	GA AAAAGGCTCTTTTAAAGACCTTACTCTTCATCACTCACTCTGTAAACAGTGCAGATGAG	300
Db	334	CACAGAGGCCATTTCAAGACTCTGTTCTTCCGGACACTCTGTGAGCAGTGCAGATGAG	393
OY	301	ATTATAGAGCTACGCTGTACCGGACGATCCAGATGAGATTTTATTCTTAATGTACC	360
Db	394	ATTATAGCTTACCATTAACGACCAATCAGAGATGTGATCTGTATTCACACGAGCA	453
OY	361	CGCTTATCATTTTGAGACCAAGAGAATATCTGTCTTATTCAAACAGACAGGCTTTATAC	420
Db	454	CGTTTACATTTTGAGAGCAGAAAGATATCTGTCTCATTCAGACAGACAGGACTTCTAC	513
OY	421	AAGCCAAAGCAGAAGTGAAGTTTGGCATGTGTACACTCTTCTCAGATTTTAAAGCTTAC	480
Db	514	AAGCCGAAGCAGGAAGTGAAGTTTCCGGTGTCTCACACTGTCTCAGATCTCAAGCTTAT	573
OY	481	AAAACTCTTTAAACATTCCTCATTAAGGACCCCAATCAATTTGATCCACAGTGGTTG	540
Db	574	AGGACATCCGGACATTTTCTATTAGGATCCGAAGTCTTAATGTATCAACAGTGGTT	633
OY	541	TCACAACAAAGTGAATCTTGAGTCAATTTCCAAACTTTTCAAGTATCTTCCATCCAAATA	600
Db	634	TCTGAAAGAGTGAATCTTGGAGTGTGTTTCCAAACTTTTCAACATCTCTAATCCAAATA	693
OY	601	CTGTGTGACTGTCTTAATTCAAAGTTCAAGTGAATGACACAGACATATTCATCATTTGAG	660
Db	694	TTTGTGTGCTGTCCATTCAGGTTCAAGTGAATGATGACAAATTTTACAGTCACTTTGAG	753
OY	661	GTTTCAGAAATGTATTAACAAATTTGAAAGTGAAGTCTTTGGACAGACCATTTATTTGCTCT	720
Db	754	GTTCTAGATATGTGTATACCAAAGTTGAAATGACCGTGACAGACCCGGCTGTATTTGTCC	813
OY	721	ATGAATTTCAAGCATTTAATATGTATACATCAACGCAAAAGTATATATATGGAAGCCAGTG	780
Db	814	TTGAAGTCTAAACAGTTAATATGCGTCGCTCATACAAAGTATATATATGGAAGCCAGTG	873
OY	781	AAAGG--AGACGTAAACGTTACATTTTATCCTTATCTTTT--GGGGAAGAGAAATAT	837
Db	874	GAAGGGAAGGTCTGTCACTTACATTTTATCCTTTGTCTTTTGGGGGAAAGAAAGAAAT	933
OY	838	ATTACAAAACATTTAAGTATTAATATGATCTGCAAACTCTCTTTTAAATGAGAAGATG	897
Db	934	ATCCAAAAGATTTTGAATATATGATTTGCAAACTTCTCTTTTGAATATATGAGATG	993
OY	898	AAAAATGTAA-----GGATCTTCAATGCACTTCTGAATACCTG	939
Db	994	AAAAAGTAAATGAATTTGAAACCGCTAACGATGTCTCGAAAGGAAGTATGAGAATGTG	1055
OY	940	GATCTATCTTCCCTCGACACGTATGAAATTTTAAACACAGTACAGAAATGAGTTACAGGT	999
Db	1054	GACCGTCTCTTCTCTGGCGCGGACAGAAATTTATCTACTGTGACAGGAATCACTACAGGT	1113

1000 ATTTCAGAAAATGTAAGCACTAATGTGTTCTTCAAGCAACATGATTTACATCATTTGAGTTT 1059  
1114 ATCTCAAAATGCGACAGACCAATGTGTTTTCAGCAACATGATTTACATTTGAGTTT 1173  
1060 TTGTGATTTACTACTGTCTTGAAGCATCTCTCAACTTCAACGCCACTGTGAAGTTAACT 1119  
1174 TTGTATTTACGACGTGTGTAACCATCTCTCAACTTCAACGCCACTGTGAAGTTAACT 1233  
1120 CGTGTGATGCGAACCACTGACTCTTGAAGAAAATAATGATGATTAACG-T 1178  
1234 CGCTGTATGCGAACCACTGACTCTTGAAGAAAATAATGATGATTAACG-T 1293  
1179 GACACAGAAAATCTACTGAGTACTGAGCGGATCTTAACAGTGGAAATCAGAAAATGGA 1238  
1294 GACCCAGGGAATAATACCAACCCGAGAG-----TCAGAGAGATCAGGAAATGGA 1344  
1239 AGCTGTTCAAAAATTAATTTATCTGTCCCAAGGAACTTTTAAGATTGAATTCCTC 1298  
1345 TTATATCCAGACCGTAATTTATCTATCCCAAAATGGAATCAATTAAGATTGAATTCCTC 1404  
1299 AATCCTGAGAGATTCAGTGAAGCTACATTTGAAGGCCATTTCTGTGTATTAAGATG 1358  
1405 AGTCAATGCCATTTCCGCGAGCTACAACTGAAGGCTTATTTCTGTATGATCAAGTTTC 1454  
1359 CATGCGACTTCAATAGTCTGTTTAAGTCTCTAGTGAACATATCACTAAATAAACAAG 1418  
1465 TGTGACACTCCACAGATATGTTCACTCTCTAGTGAACATATCACTAAATAAACAAG 1524  
1419 AGATGAAAATTAAGAGTGGAGTCGCTTTGAGTGTGTGTAGTGGCAACAAAGAT 1478  
1525 AGATGAAAATTAAGAGTGGAGTCGCTTTGAGTGTGTGTAGTGGCAACAAAGAT 1584  
1479 GAAAGAGTGAAGCTATAGTGAATTCAGAGGAGAGTGTGTGTGTAGTGAAGAAAACAAG 1538  
1585 CAAAGAGTGAAGCTATAGTGAATTCAGAGGAGAGTGTGTGTGTAGTGAAGAAAACAAG 1644  
1539 TTCAACAATGTTCTCTTAAACACAGAAAATTTCTGAGTCCAAAAGCTGTGTATTTGT 1598  
1645 TTCAACAATGTTCTCTTAAACACAGAGGCTTCCGAGGCTCCAAAAGCTGTGTATTTGT 1704  
1599 GTATTAATTTGAAGATGAGTGGGAAATTAATGAATGTTCTTAAATTCCTGTTCAGCT 1658  
1705 ATATTAATTTGAAGATGAGTGGGAAATTAATGAATGTTCTTAAATTCCTGTTCAGCT 1764  
1659 TGTATTAATTTGAAGATGAGTGGGAAATTAATGAATGTTCTTAAATTCCTGTTCAGCT 1718  
1765 TGTATTAATTTGAAGATGAGTGGGAAATTAATGAATGTTCTTAAATTCCTGTTCAGCT 1824  
1719 AGTCTCTTGAAGATCTCTGTGAACAGAGCTGATCCATGATGAGTGTGTAGTGTGA 1778  
1825 GGTCTCTTGAAGATCTCTGTGAACAGAGCTGATCCCTGTGTGTGTGTGTGTGA 1884  
1779 CAAAAGTGTGAATCTGATGAATGCTCTTAATGATTAATCAATGAGAAAATGTGTTCATGA 1838  
1885 CAAAAGTGTGAATCTGATGAATGCTCTTAATGATTAATCAATGAGAAAATGTGTTCATGA 1944  
1839 GTTGAATCTTAAATCAAGAGATTAATTTAGAGCATTTTCAATGATTTCTTTTGAAGTGT 1898  
1945 ACTGGAATCTTAAATCAAGAGATTAATTTAGAGCATTTTCAATGATTTCTTTTGAAGTGT 2004  
1899 TCAGGAATGAGGATCTGTGGATTTGAAGATGCAAACTCCGAGAGATTAATTTGAAGTGT 1958  
2005 TCAGGAATGAGGATCTGTGGATTTGAAGATGCAAACTCCGAGAGATTAATTTGAAGTGT 2064  
1959 TGTATTAATCAATGAGATTAATGCTGAGAGGTTTAATGAGAGAAAATGAGAGCATATTTGT 2018  
2065 AGTCAATGATCTGAAGAGATTAATGAGAGGTTTAATGAGAGAAAATGAGAGCATATTTGT 2124  
2019 AGATTAATCAATGATCTTTTGT 2078  
2125 TGTATTAATCAATGATCTTTTGT 2184

2079 TTGATTTGCTGACAGCAACATGAGTTTACAGAGTTTACAGAAATTTGAAGTAACTGT 2138  
2185 ATGATTTGCTGACAGCAACATGAGTTTACAGAGTTTACAGAAATTTGAAGTAACTGT 2244  
2139 ACCGATTTCAATCACTTCTTGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2198  
2245 ACCGATTTCAATCACTTCTTGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2304  
2199 TGACATTAACAACTACTGAGTGTGAGCTCCAAAGCTTCCAACTTTTCAATTTTGTGA 2258  
2305 TGGGTTAACAACTACTGAGTGTGAGCTCCAAAGCTTCCAACTTTTCAATTTTGTGA 2364  
2259 TCTTCCCTACTCTGTATTCAGAGGTGAAGATTTGCTTTGAAATTAATTAATTAATTA 2318  
2365 TCTTCCCTACTCTGTATTCAGAGGTGAAGATTTGCTTTGAAATTAATTAATTAATTA 2424  
2219 TTGAAAGATGCACTGAGTGTGAAGTAACTTTGAAGAAAATGAGAGCAATTTGATATCT 2378  
2425 TTGAAAGATGCACTGAGTGTGAAGTAACTTTGAGTAAAGTAAAGTAAAGTAAAGTAA 2484  
2379 AATGACTTCAAGTGAATTAATTAATGCAAGGCAACAGAGCAACCTTCTGTGTCCAGTGA 2438  
2485 AATGACTTCAAGTGAATTAATTAATGCAAGGCAACAGAGCAACCTTCTGTGTCCAGTGA 2544  
2439 GATGAGGCAACTGTTCTTTTCCATCAGGCAACATCTGAGAGAAATTCCTATCAC 2498  
2545 TAATGAGGCAACTGTTCTTTTCCATCAGGCAACATCTGAGAGAAATTCCTATCAC 2604  
2499 AGTCAAGCTCTTTCACCCACTGCTTGTATGCTGTGACCAAGATTTTGAAGAGC 2558  
2605 AGTCAAGCTCTTTCACCCACTGCTTGTATGCTGTGACCAAGATTTTGAAGAGC 2664  
2559 TGAAGAAATGAAAATCAATTAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 2618  
2665 TGAAGAAATGAAAATCAATTAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 2724  
2619 ACAAGATGACCTGAAAATTTGAAGTGTCTGATTTCTCTTAATTAATTAATTAATTAAT 2678  
2725 AGAAGATGACCTGAAAATTTGAAGTGTCTGATTTCTCTTAATTAATTAATTAATTAAT 2784  
2679 AAGAGTCAATCACTGCAATTTGAAGTGTCTGATTTCTCTTAATTAATTAATTAATTAAT 2738  
2785 AAGAGTCAATCACTGCAATTTGAAGTGTCTGATTTCTCTTAATTAATTAATTAATTAAT 2844  
2739 AATGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 2798  
2845 AATGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 2904  
2799 TTAATTTTGAATTAATGATCTAATAAAGAAACAATGAGATTAATTTGAAGAAAAGC 2858  
2905 TTAATTTTGAATTAATGATCTAATAAAGAAACAATGAGATTAATTTGAAGAAAAGC 2964  
2859 TCTTCAATTTATGAGGCAAGTAACTTCTTATCAAGAGGAAATGAGTGT 2918  
2965 CTTTCAATTTATGAGGCAAGTAACTTCTTATCAAGAGGAAATGAGTGT 3024  
2919 TTTCAAGTGTGAGGAAATTAATGACCTTCTGAGAGCACTTGTGTGAGTGTGTGTGTGT 2978  
3025 CTTCAAGTGTGAGGAAATTAATGACCTTCTGAGAGCACTTGTGTGAGTGTGTGTGTGT 3084  
2979 AAGATGTTCTTGAAGCCGATCTTAATGATTAATTAATTAATTAATTAATTAATTAATTA 3038  
3085 AAGATGTTCTTGAAGCCGATCTTAATGATTAATTAATTAATTAATTAATTAATTAATTA 3144  
3145 AATGATGTTCTTGAAGCCGATCTTAATGATTAATTAATTAATTAATTAATTAATTAATTA 3204  
3099 GATTTATGAGGAGTGTGAAGTGTGAAGTGTGAAGTGTGAAGTGTGAAGTGTGAAGTGT 3158  
3205 GATTTATGAGGAGTGTGAAGTGTGAAGTGTGAAGTGTGAAGTGTGAAGTGTGAAGTGT 3264  
3159 AATTTCTCTCTGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3218



```

Db      3265 GACTTGTCTCGGAGTACAAAAAGTATAGCCTAATATGATGACAAAGCTCATCA 3324
Qy      3219 TTTTGGAGTCTGAATTCAGTATAGAGAAATTCAGCAATTAATCTAGCCCTTAAC 3278
Db      3325 GTTTTGGATTGTAATTCAGCAGAGAAATTCGACAAATTAATCTTAGCAATTAATTC 3384
Qy      3279 TTAATGATTCATCATAGTGGGAGTCTTAAGCGAAGAAAGCTTGAATATAGTGA 3338
Db      3385 CTAATGCTCTGTCACAGTCGGGAGCCCTTAAGCAGAGAGGCTTTGAATCTTGATGA 3444
Qy      3339 GAGAGAGAAACAAGAAAGTGGCATATTCGGGTGATCAGAGTCCAACTTTCTGA 3398
Db      3445 GGGATCAGAAAAGAGAGAGACACAGATTCGGTGTATCAGAGACCTGCTCTCTG 3504
Qy      3399 CTCCTGGCAGCAGCTCCCTGGATATTAAGTTCAGCTTCATGCTCTCACACT 3458
Db      3505 GTCTGGAGCCGCGCTCCGTGACATCGAAATCGCTTATGACCTGCTGGCAGAC 3564
Qy      3459 CTTAATTTTCAAGATCTTGAAGGAATCCCAATTAAGAGTGGCTTAAGCAGCAAGAA 3518
Db      3565 TCTGCA---CCATGTCTGAGGAAATCCGGTTATGAGTGGCTCATCCAGCAAGAA 3621
Qy      3519 TAGCTTGGGTGTTTTCATCTAATCAGATACCACTGGCTTTAAGGCTCTGTGA 3578
Db      3622 CAGCTGGAGGTTTGTATCCACTCAGATATCTGTTGGCTGAGGCAATTAATCTGA 3681
Qy      3579 ATTGGAGCCCTAATGAATACAAAGAAACAATATCCAAAGTCCGTAAGCGGCGCTAG 3638
Db      3682 ATTTTACGCCCTTGTCACAAAGAAATACAGTATCAACGTGACCGTACAGGCGCCG 3741
Qy      3639 CTCACCAAGTCTT----- 3651
Db      3742 CATCCCAATCATACATTCGCAATGACTCCAGAACCTTTCTCTTCCACAGCA 3801
Qy      3652 ----CTTGCTGTGTGACAGCCAAATGCAATTAATTTCCGAATGTTTGGATTTGC 3707
Db      3802 AGAGCTTACGCGCTGATCCATCCATCCGTTAATGTTTCTGCGCATGCTCGGAAATTC 3861
Qy      3708 TATTTGTGAGTCAATGTTTATATATATGAAAGCTTGGCTTCTTGAAGAGCAAG 3767
Db      3862 TATTTGCGCAATTAATGTTGACTATTAACGTAAAGGTTCTTCTTAAGACGAG 3921
Qy      3768 ATCTATCCAAATTAAGAAAGCTTTGATTTAGTGTGCTGTAAGAAATTAAGATGA 3827
Db      3922 ATCTACCGAAATTAAGAAAGTTTGTATTAAGCTGATTTGA---AATATGAGAGCA 3978
Qy      3828 TCTCATCATGTGTGATTAATGTGTGTAACAAGCTTTGCGGCCGCGGTAGAGTGCAT 3887
Db      3979 CATGTGACCTGAAATCTGAATGTGTGCAACAAGTCACTTGGGTTCAAGAGACAGGCAT 4038
Qy      3888 GGCTCTTAATGAAGTAACTAATTAATGAGCTTTAGTCCCTTCAAGAGCAATTTCT 3947
Db      4039 GGTCTTAATGAAGTAACTTCTCAGTGGCTTATGATGATCTTCAATTAATTCCT 4098
Qy      3948 GAGCGAGACAGTGAAGAAAGTGAATATGATCATGSAAACTCAACTTATTTAGATTC 4007
Db      4099 GAGTGAAGACCTGAAAGAAAGTGAATATATTAACGGAAATCAACTTATTTAGATTC 4158
Qy      4008 TGTAAATGAACCCAGTTTGTGTAAATATCTCTGCTGAGAACTTTAAAGTTTCAA 4067
Db      4159 TGTAAATGAATCCAGTTTGTGTAAATATCCCACTGTGAGACTAACAAAGTTTCAA 4218
Qy      4068 TACCCAGATGCTTCAAGTCCATAGTGAATTAATGAGCCAGAGAGCAGCGGTGAG 4127
Db      4219 TATTTGAGATGTTCTGATCTGTATGAGACTAATAAGAACAAAGAGCAGCGAGTGG 4278
Qy      4128 AAGTTCAACTCTGAAGTGAAGCTCTCTCTGTGACTTTTGAAGTGAAGTCAAGGCTG 4187
Db      4279 AAGCTCAACACCCAGGTGAAGTGTCTCTGTGTTACT---CAGTCCAGAACCAACTG 4335
Qy      4188 CCGTCTTGTGAGATGAGCTTCAAGCTTCCATCATCACTTCAAGTATTTTATTTT 4247

```

```

Db      4336 CAATGACACACAGCAGGAGCCAGCAAGACTCCCTTGACGTTCTTGACCTTCTGCTT 4395
Qy      4248 CTGTTCAAGCTTCTGATCTTATATGAACTTTAGCTGTGA 4287
Db      4336 TTGCTCCGCTCTTCTGTAATTTTGACAAATGATTAAGA 4435

RESULT 3
AL554092
LOCUS
DEFINITION
AL554092 1201 bp mRNA linear EST 31-MAY-2003
clone CS0D1081YB18 5-PRIME, mRNA sequence.
ACCESSION
AL554092
VERSION
AL554092.2 GI:31275905
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE
1 (bases 1 to 1201)
Li, W.B., Gruber, C., Jessup, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
JOURNAL
On Feb 15, 2001 this sequence version replaced gi:12894544.
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 3578.f for
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0D1081YB18&cluster=3578.f. Contact :
Feng Liang Email : fliang@life.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0D1081YB18

FEATURES
source
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1081YB18"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="First strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT
359 a 238 c 236 g 293 t 75 others
ORIGIN

Query Match 20.8%; Score 893; DB 9; Length 1201;
Best Local Similarity 93.9%; Pred. No. 2,3e-187;
Matches 947; Conservative 39; Mismatches 14; Indels 9; Gaps 6;

1 ATGCAAGGCGCCAGCGCTCTGACCGCGCCCACTCTCGGTGTGACCGCGCGTG 60
Db      92 ATGCAAGGCGCCAGCGCTCTGACCGCGCCCACTCTCGGTGTGACCGCGCGTG 151
Qy      61 GCCGTGCTCCCGGCGCTCGGTCTGTGTGACAGCCCGAGGATCATCAGGCCGAGAGA 120
Db      152 GCCGTGCTCCCGGCGCTCGGTCTGTGTGACAGCCCGAGGATCATCAGGCCGAGAGA 211
Qy      121 AATGTGACTATTTGGGTGAGACTTCTGGAACAATGCTTCAAGTGAATGAGCG 180
Db      212 AATGTGACTATTTGGGTGAGACTTCTGGAAMAATGCTTCAAGTGAATGAGCG 271
Qy      181 GAGTGTCAAGAGACATCAACTCACTGTCTGTCTGCGAAGAGAGAGTCTT 240
Db      272 GAGTGTCAAGAGACATCAAAATTCATCTCTGTGTCGAGAGAGAGTCTT 331
Qy      241 GAAAAAGGCTTTTAAGACACTTCTTCATCACTCACTGAAACAGTGCAGATGAG 300

```

Db 332 GAAAAAGCTCTTTTAAAMACTTAMCTTCCATMAMTACTGAAACATGACATGATGAG 391  
 Qy 301 ATTATGAGCTACGCTTAACCGGACGTACCCAGAGTAGATTTATTTCTTATAGTACC 360  
 Db 392 ATTATGAGCTACGCTTAACCGGACGTACCCAGAGTAGATTTATTTCTTATAGTACC 451  
 Qy 361 CGCTTATGAGCTACGCTTAACCGGACGTACCCAGAGTAGATTTATTTCTTATAGTACC 420  
 Db 452 CGCTTATGAGCTACGCTTAACCGGACGTACCCAGAGTAGATTTATTTCTTATAGTACC 511  
 Qy 421 AAGCCAAAGCAAGAGTGAAGTGGCTGCTTGTATGACCTTCTCAGATTTTAAAGCTTAC 480  
 Db 512 AAGCCAAAGCAAGAGTGAAGTGGCTGCTTGTATGACCTTCTCAGATTTTAAAGCTTAC 571  
 Qy 481 AAGCCCTTTTAAAGTGAAGTGGCTGCTTGTATGACCTTCTCAGATTTTAAAGCTTAC 540  
 Db 572 AAGCCCTTTTAAAGTGAAGTGGCTGCTTGTATGACCTTCTCAGATTTTAAAGCTTAC 631  
 Qy 541 TCACAAACAAAGTGAAGTGGCTGCTTGTATGACCTTCTCAGATTTTAAAGCTTAC 600  
 Db 632 TCACAAACAAAGTGAAGTGGCTGCTTGTATGACCTTCTCAGATTTTAAAGCTTAC 691  
 Qy 601 CTGTGAGCTGCTTATGACCTTGTATGACCTTGTATGACCTTGTATGACCTTGTATGAC 660  
 Db 692 CTGTGAGCTGCTTATGACCTTGTATGACCTTGTATGACCTTGTATGACCTTGTATGAC 751  
 Qy 661 GTTTCAGATTAATGATTAACCAAAATTTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 720  
 Db 752 GTTTCAGATTAATGATTAACCAAAATTTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 811  
 Qy 721 ATGAATTTCAAGCAATTTAATGATTAACCAAAATTTGAGTGAAGTGAAGTGAAGTGAAGTGA 780  
 Db 812 ATGAATTTCAAGCAATTTAATGATTAACCAAAATTTGAGTGAAGTGAAGTGAAGTGAAGTGA 870  
 Qy 781 AAAGAGAGCTGACCTTACATTTTACCTTATCTTTTGGGAAAGAAAGAAATTT 840  
 Db 871 AAAGAGAGCTGACCTTACATTTTACCTTATCTTTTGGGAAAGAAAGAAATTT 930  
 Qy 841 ACAAACAATTTAATGATTAACCAAAATTTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 900  
 Db 931 ACAAACAATTTAATGATTAACCAAAATTTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 990  
 Qy 901 AATGATGATTTTCAAGCAATTTAATGATTAACCAAAATTTGAGTGAAGTGAAGTGAAGTGA 960  
 Db 991 AATGATGATTTTCAAGCAATTTAATGATTAACCAAAATTTGAGTGAAGTGAAGTGAAGTGA 1045  
 Qy 961 GTAGAAATTTTCAAGCAATTTAATGATTAACCAAAATTTGAGTGAAGTGAAGTGAAGTGA 1009  
 Db 1046 GTAGAAATTTTCAAGCAATTTAATGATTAACCAAAATTTGAGTGAAGTGAAGTGAAGTGA 1091  
 RESULT 4  
 LOCUS BX391819/c 941 bp mRNA linear EST 13-MAY-2003  
 DEFINITION BX391819 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
 clone CS001081YB18 3-PRIME, mRNA sequence.  
 ACCESSION BX391819  
 VERSION BX391819.1 GI:30627404  
 KEYWORDS EST.  
 ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 941)  
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished  
 COMMENT Genoscope - Centre National de Sequencage  
 BP 191 91006 Evry cedex - France  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 Library was constructed by Life Technologies, a division of  
 Invitrogen. This sequence belongs to sequence cluster 3578.f For

more information about this cluster, see  
<http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS001081YB18>  
 Contact : Feng Liang Email : fliang@lifetech.com url :  
<http://fulllength.invitrogen.com/InvitrogenCorporation1600>  
 Faraday Avenue Genoscope sequence ID : CS001081YB18\_CS02740\_1.  
 Location/Qualifiers  
 1. 941  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS001081YB18"  
 /library="PLACENTA COT 25-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."  
 BASE COUNT 271 a 214 c 181 g 266 t 9 others  
 ORIGIN  
 Query Match 20.4%; Score 872.8; DB 13; Length 941;  
 Best Local Similarity 96.7%; Pred. No. 6.8e-183;  
 Matches 909; Conservative 0; Mismatches 28; Indels 3; Gaps 2;  
 Db 2820 TAAAAAGAAACAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2879  
 Qy 939 TAAAGAAACAGCTGACAGATTTTGAAGAAAGAAAGCTCTTNTCATTTAGAGGCAAG 880  
 Db 2880 TTAACAGAGAACTTCTTATGAGAGGAGTGGCTCTTGAAGTGAAGTGAAGTGAAGTGA 2939  
 Qy 879 GTACAGAGAGACTTCTTATGAGAGGAGTGGCTCTTGAAGTGAAGTGAAGTGAAGTGA 821  
 Db 2940 TGAACCTTGGAGAGCACTTGTGAGCTTTTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2999  
 Qy 820 TGAACCTTGGAGAGCACTTGTGAGCTTTTGAAGTGAAGTGAAGTGAAGTGAAGTGA 761  
 Db 3000 TCCTTACATGATTTGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3059  
 Qy 760 TCCTTACATGATTTGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 701  
 Db 3060 TCAAGAAATCAAGCGTGAATTTGGATCCAGAGAGTGAATTCATGAGCTTCAAG 3119  
 Qy 700 TCAAGAAATCAAGCGTGAATTTGGATCCAGAGAGTGAATTCATGAGCTTCAAG 641  
 Db 3120 TGGCAATTAAGTCCAGTACCTTACAGCTTATGATTTTGGAGTGAAGTGAAGTGAAGTGA 3179  
 Qy 640 TGGCAATTAAGTCCAGTACCTTACAGCTTATGATTTTGGAGTGAAGTGAAGTGAAGTGA 581  
 Db 3180 AAAGTACAGCTTACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 3239  
 Qy 580 AAAGTACAGCTTACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 521  
 Db 3240 TAAGAGATTTCAAGCAATTTAATGATTTTGGAGTGAAGTGAAGTGAAGTGAAGTGA 3299  
 Qy 520 TAAGAGATTTCAAGCAATTTAATGATTTTGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 461  
 Db 3300 GAGTCCTAAGCGAAGAGAGCTTGAATAGCTGAGTGAAGGAGCAAGAAAGGTGG 3359  
 Qy 460 GAGTCCTAAGCGAAGAGAGCTTGAATAGCTGAGTGAAGGAGCAAGAAAGGTGG 401  
 Db 3360 CATGCAATTTGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 3419  
 Qy 400 CATGCAATTTGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 341  
 Db 3420 GGAATTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 3479  
 Qy 340 GGAATTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 281  
 Db 3480 GGAATTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 3539  
 Qy 280 GGAATTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 221

QY 3540 TACTGAGATACCACTGCTGCTTTAAAGCTCTGTCTGAATTTGACGCCCTTAATGATNC 3599  
| | | | |  
Db 220 TACTGAGATACCACTGCTGCTTTAAAGCTCTGTCTGAATTTGACGCCCTTAATGATNC 161  
| | | | |  
QY 3600 AGAAGAGCAAAATATATCAAGTACCGGTGACGGGGCTTACTACCAAGTCTCTGCTCT 3659  
| | | | |  
Db 160 AGAAGAGCAAAATATATCAAGTACCGGTGACGGGGCTTACTACCAAGTCTCTGCTCT 101  
| | | | |  
QY 3660 GGTACAGCCCAATGACGATTATATTTCCGCAATGGTTTGGATTGTATTTGTCAGCT 3719  
| | | | |  
Db 100 GGTACAGCCCAATGACGATTATATTTCCGCAATGGTTTGGATTGTATTTGTCAGCT 41  
| | | | |  
QY 3720 CAATG--TTGTATATATATGTAAGGCTTTGGGCTCTTCTA 3757  
| | | | |  
Db 40 CAATGTTTGTATATATATGTAAGGCTTTGGGCTCTTCTA 1  
| | | | |  
RESULT 5  
BUS17328 879 bp mRNA linear EST 12-SEP-2002  
LOCUS BUS17328  
DEFINITION AGENCOURT 10164521 NIH\_MGC\_71 Homo sapiens cDNA clone IMAGE:6514905  
5', mRNA sequence.  
ACCESSION BUS17328  
VERSION BUS17328.1 GI:22824854  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
1 (bases 1 to 879)  
NIH-MGC http://mgi.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: cge@bbs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLNL14090 row: P column: 10  
High quality sequence stop: 663.  
Location/Qualifiers  
1. 879  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6514905"  
/tissue\_type="leiomyosarcoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_71"  
/note="Organ: uterus; Vector: pCMV-Sport6; site: 1; NotI;  
site: 2; SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 2.1 kb." 1 others  
BASE COUNT 248 a 171 c 206 g 253 t  
ORIGIN  
Query Match 16.6%; Score 712.4; DB 13; Length 879;  
Best Local Similarity 91.8%; Pred. No. 3e-147;  
Matches 807; Conservative 0; Mismatches 17; Indels 55; Gaps 3;  
QY 3207 AAGAGCTATCCATTTTGGAGTCTGCAATTCAGTAGAGAAATTCAGCAATTAATCTT 3266  
| | | | |  
Db 1 AAGAGCTATCCATTTTGGAGTCTGCAATTCAGTAGAGAAATTCAGCAATTAATCTTCT 60  
| | | | |  
QY 3267 ACCCTTATATATGATGATGATGATGAGGAGTCTCTAAAGCAAGAAAGCTTTGAA 3326  
| | | | |  
Db 61 ACCCTTATATATGATGATGATGATGAGGAGTCTCTAAAGCAAGAAAGCTTTGAA 120  
| | | | |  
QY 3327 TATGCTGATGAGAGCAAGCAAGAGGTGATGCAATTCGTGGTGTCAATCAAGATC 3386  
| | | | |

Db 121 TATGCTGATGAGAGCAAGCAAGAGGTGATGCAATTCGTGGTGTCAATCAAGATC 180  
| | | | |  
QY 3387 CAACCTTTCTGACTCTCTGAGCAGCAGCTCCCTGATATTAAGTTGACGCTTATGACT 3446  
| | | | |  
Db 181 CAACCTTTCTGACTCTCTGAGCAGCAGCTCCCTGATATTAAGTTGACGCTTATGACT 240  
| | | | |  
QY 3447 GCTCTCACACTTCTTACAAATTTGACATCTTGAAGGAATCCCAATTAAGGTGCTAAG 3506  
| | | | |  
Db 241 GCTCTCACACTTCTTACAAATTTGACATCTTGAAGGAATCCCAATTAAGGTGCTAAG 300  
| | | | |  
QY 3507 CAGGCAAAAGAAATTAATCTTGGGTGTTTGGATCTCTCAGATATACCACTGCTTTAA 3566  
| | | | |  
Db 301 CAGGCAAAAGAAATTAATCTTGGGTGTTTGGATCTCTCAGATATACCACTGCTTTAA 360  
| | | | |  
QY 3567 GCTCTCTGATTAATTTGAGCCCTTATGATATACAGAAAGCAAAATATCAAGTACGCT 3626  
| | | | |  
Db 361 GCTCTCTGATTAATTTGAGCCCTTATGATATACAGAAAGCAAAATATCAAGTACGCT 420  
| | | | |  
QY 3627 GACGGGGGCTTACCTGACCAAGTCTT----- 3651  
| | | | |  
Db 421 GACGGGGGCTTACCTGACCAAGTCTTAAAGTTCTGATGACACACCAACCGCTTACT 480  
| | | | |  
QY 3652 -----CTTGTGTGTGTACAGCAATGCGATTTATTTCCGCAAAATGG 3695  
| | | | |  
Db 481 CTTTACAGACAGAGCTTGT 540  
| | | | |  
QY 3696 TTTTGATTTGCTATTTGTGACCTGATGTTGATATATGTAAGGCTTCCGGCTTC 3755  
| | | | |  
Db 541 TTTTGATTTGCTATTTGTGACCTGATGTTGATATATGTAAGGCTTCCGGCTTC 600  
| | | | |  
QY 3756 TAGAAGAGCAAGATCTATCCAAATCAAGAGCCTTTGATTTAGTGTGCTGTAAAGA 3815  
| | | | |  
Db 601 TAGAAGAGCAAGATCTATCCAAATCAAGAGCCTTTGATTTAGTGTGCTGTAAAGA 660  
| | | | |  
QY 3816 AATATAAGATGATCTCAATCATGTGATTTGATGTGTGTGTGTGTGTGTGTGTGTGTGT 3875  
| | | | |  
Db 661 AATATAAGATGATCTCAATCATGTGATTTGATGTGTGTGTGTGTGTGTGTGTGTGTGT 720  
| | | | |  
QY 3876 TAGAGTGTGATGCTCTTATGAGATTAAGTAACTTAAGTGTGTGTGTGTGTGTGTGTGTGTGT 3935  
| | | | |  
Db 721 TAGAGTGTGATGCTCTTATGAGATTAAGTAACTTAAGTGTGTGTGTGTGTGTGTGTGTGTGT 780  
| | | | |  
QY 3936 AGCAATTTCT--CTGAGCAGACAGTGAAGAAAGTGAATATGATCATGGAATCAACC 3994  
| | | | |  
Db 781 AGCAATTTCTCTGAGCAGACAGTGAAGAAAGTGAATATGATCATGGAATCAACC 840  
| | | | |  
QY 3995 TCT--ATTAGATTTCTGTAATGAACCCAGTTTGTG 4030  
| | | | |  
Db 841 CCTTATTTTGAATTTCTGTAATGAACCCAGTTTGTG 879  
| | | | |  
RESULT 6  
BX386457 978 bp mRNA linear EST 08-MAY-2003  
LOCUS BX386457 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
DEFINITION clone CS01081YB18 5-PRIME, mRNA sequence.  
ACCESSION BX386457  
VERSION BX386457.1 GI:30457430  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
1 (bases 1 to 978)  
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 3578.f for

more information about this cluster, see  
<http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS1A1021ZB100P1&cluster=3578.f>. Contact :  
 Feng Liang Email : [fliang@lifetech.com](mailto:fliang@lifetech.com) URL : <http://fulllength.invitrogen.com/> Invitrogen Corporation 1600  
 Paraday Avenue Genoscope sequence ID : CS1A1021ZB100P1.

## FEATURES

## Source

1..978  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0D1081YB18"  
 /issue\_type="PLACENTA COT 25-NORMALIZED"  
 /clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."  
 BASE COUNT 374 a 151 c 179 g 220 t  
 ORIGIN

Query Match 16.6%; Score 710.8; DB 13; Length 978;  
 Best Local Similarity 84.6%; Pred. No. 6.9e-147;  
 Matches 758; Conservative 46; Mismatches 90; Indels 2; Gaps 2;

7 GGCCACGCGCTCTGACCGCCGCCACCTCTGCGGTGACGCGCGCGTGGCGTG 66  
 84 GGGCCACGCTCTGACCGCCGCCACCTCTGCGGTGACGCGCGCGTGGCGTG 143  
 67 GCTCCCGGCGCTCTGCGGTGACGCGCGCGCGCGTGGCGTG 126  
 144 GCTCCCGGCGCTCTGCGGTGACGCGCGCGCGCGTGGCGTG 203  
 127 ACTATGGGGGAGAGCTTCTGGAACACCTGCTTCAAGGTGATGTAAGCGGAGTG 186  
 204 ACMAATGGGGGAGAGCTTCTGGAACACCTGCTTCAAGGTGATGTAAGCGGAGTG 262  
 187 CTCAAGACAGCATCAACCTGCTCTGCTGCTGGAAGAGAGAGTCTTTGAAAAA 246  
 263 CTCAAGACAGCATCAACCTGCTCTGCTGCTGGAAGAGAGAGTCTTTGAAAAA 322  
 247 GGGCTTTTAAAGACCTTACTCTTCCATCCTGCTGGAAGAGAGTGAATTTAT 306  
 323 AGATCTTTTAAAGACCTTACTCTTCCATCCTGCTGGAAGAGAGTGAATTTAT 382  
 307 GAGTACGCTGTAACCGGACGTAACCGAGATGAGTTTATCTCTAATAGTACCGCTTA 366  
 383 AARCMAGTGTAAAGAGAGTACCGAGATGAGTTTATCTCTAATAGTACCGCTTA 442  
 367 TCATTGAGCAAGAGAGATATCTGCTTCAATTCAAAAGAGAGCTTTAAGAGCA 426  
 443 AAAATTGAAACCAAAAAATTAAGAGATTAAGAGAGAGAGATTAAGAGCA 502  
 427 AAGCAAGAGTGAAGTTCGATGTTTACCTCTTCTGAGATTTTAAAGCTTTACAAAC 486  
 503 AAGCAAGAGTGAAGTTCGATGTTTACCTCTTCTGAGATTTTAAAGCTTTACAAAC 562  
 487 TCTTTAAACATTTCTATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 546  
 563 ACTTTAAACATTTCTATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 622  
 547 CAAAGTGAATCTGAGTCAATTTCCAAACCTTTGAGTATCTTCCATCCAAATCTGGT 606  
 623 CAAAGTGAATCTGAGTCAATTTCCAAACCTTTGAGTATCTTCCATCCAAATCTGGT 682  
 607 GACTGCTATTCAGATTCAGATGATGATGATGATGATGATGATGATGATGATGAT 666  
 683 GACTGCTATTCAGATTCAGATGATGATGATGATGATGATGATGATGATGATGAT 742  
 667 GAATATGATTAACCAAAATTTGAAGTGAATTTGACAGACACATTAATTTGTTATGAT 726  
 743 GAATATGATTAACCAAAATTTGAAGTGAATTTGACAGACACATTAATTTGTTATGAT 802

Qy 727 TCTAAGCATTTAATGTTACCATCAGCGAAGATATACATATGAGAGACCGATGAAGA 786  
 Db 803 TCTAAGCATTTAATGTTACCATCAGCGAAGATATACATATGAGAGAGACCGATGAAGA 862  
 Qy 787 GAGTACGCTTACATTTTACCTTTTACCTTTTGGG -AAAGAGAAAAATTTACAA 845  
 Db 863 GACTTAACCTTCAATTTTACCTTTTACCTTTTGGG -AAAGAGAAAAATTTACAA 922  
 Qy 846 AACATTAAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 901  
 Db 923 AACATTAAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 978

RESULT 7  
 B0028040/c  
 LOCUS  
 DEFINITION  
 UI-H-CO0-ath-c-08-0-UI-81 NCI\_CGAP\_Sub9 Homo sapiens cDNA clone  
 IMAGE:3106526 3, mRNA sequence.  
 ACCESSION  
 B0028040  
 B0028040.1 GI:19763319  
 KEYWORDS  
 EST.  
 SOURCE  
 Homo sapiens (human)  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 667)  
 NCI\_CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 Unpublished  
 Contact: Robert Straube, Ph.D.  
 Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
 Tissue Procurement: Dr. Jose Mercuende  
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Clone distribution information can be found  
 through the I.M.A.G.E. Consortium/HLN at: <http://image.llnl.gov>  
 Seq primer: M13 FORWARD  
 POLY-A=yes.

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

Unpublished  
 Contact: Robert Straube, Ph.D.  
 Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
 Tissue Procurement: Dr. Jose Mercuende  
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Clone distribution information can be found  
 through the I.M.A.G.E. Consortium/HLN at: <http://image.llnl.gov>  
 Seq primer: M13 FORWARD  
 POLY-A=yes.

## FEATURES

## Source

## Location/Qualifiers

1..667  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3106526"  
 /issue\_type="mixed"  
 /dev\_stage="mixed"  
 /lab\_host="DH10B (Life Technologies)"  
 /note="Vector: pT7T3-Pac (Pharmacia) with a modified  
 polylinker; Site 1: EcoR I; Site 2: Not I; tissues:  
 Chondrocytes with Crohn's disease, Chondrocytes with  
 ulcerative colitis, Cervix thymus, Cervix, Cervical  
 adenocarcinoma, Bladder carcinoma, Brain oligodendrocyte,  
 NCI CGAP Sub9 is a subcloned cDNA library constructed  
 according to Bonaldi, Lemon and Soares, Genome Research,  
 6:791-806, 1996. First strand cDNA synthesis was primed  
 with an oligo-dT primer containing a Not I site. Double  
 stranded cDNA was ligated to an EcoR I adaptor, digested  
 with Not I, and cloned directionally into pT7T3-Pac  
 vector. The oligonucleotide used to prime the synthesis of  
 first-strand cDNA contains a library tag sequence that is  
 located between the Not I site and the (dT)18 tail. The  
 sequence tags for this library are GTC, AACG, GGGCC,  
 GGAAG, TAGC, TAAAC, ATGG, AGACA, ATCAC. For additional  
 information, contact: Bento Soares, [bento-soares@iowa.edu](mailto:bento-soares@iowa.edu)  
 TAG LIB=UI-H-CO0  
 TAG TISSUE=Cervical Adenocarcinoma  
 TAG SEQ=CGAGG  
 BASE COUNT 194 a 113 c 139 g 221 t  
 ORIGIN









Db	421	GGTGGCCGCTCTGGAGGATGAGACTTCAGGCTCCATCATCACTTCAGTCAATTTT	480
Db	423	ATTTCCTGTTCAAGCTCTGTACTTATGACACTTGGCTGTGA	4287
Db	481	ATTTCCTGTTCAAGCTCTGTACTTATGACACTTGGCTGTGA	525
RESULT 12			
LOCUS	B6E13975		
DEFINITION	B6E13975 RIKEN full-length enriched, 0 day neonate head Mus musculus cDNA clone 4832419101 5', mRNA sequence.		
ACCESSION	B6E13975		
VERSION	B6E13975.1	GI:16454440	
KEYWORDS	EST.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Komno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasakita,D., Shibata,K., Shingawa,A., Shirai,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M., and Hayashizaki,Y.		
TITLE	RIKEN Mouse ESTs (Arakawa,T., et al. 2001)		
JOURNAL	Unpublished		
COMMENT	Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@gscc.riken.go.jp/ URL: http://genome.gsc.riken.go.jp/ Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Komno,H., Okazaki,Y., Muramatsu,M., and Hayashizaki,Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000) Wagui,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsunura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A. and Hayashizaki,Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000) Komno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y., and Hayashizaki,Y. Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001) Kondo,S., Shingawa,A., Saito,T., Kiyosawa,H., Yamataka,I., Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and Hayashizaki,Y. Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001) Please visit our web site (http://genome.gsc.riken.go.jp) for further details. e mouse tissues. Location/Qualifiers 1..648 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="4832419101" /sex="mixed" /tissue_type="head" /dev_stage="0 day neonate"		
FEATURES			
SOURCE			

```

/lab host="DHI0B"
/closure_lib="RIKEN full-length enriched, 0 day neonate
head"_lib="RIKEN full-length enriched, 0 day neonate
/head"-Site_1: SalI; Site_2: BamH; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN, Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGGAGAGAAGATCCAAAGACCTTTTTCCTTTTTTTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot = 10.0 and subtraction to Rot = 100.0. Second
strand cDNA was prepared with the primer adaptor of
sequence [5' GAGGAGAGATTCCTCGATTAAATAAATATCCTCCCCCCCC
3']. cDNA was cloned into the XhoI and BamH sites.
Vector: a modified Bluescript KS(+) after bulk excision
from Lambda Phage I."

BASE COUNT      195 a      118 c      140 g      194 t      1 others
ORIGIN

Query Match          11.7%; Score 501.6; DB 10; Length 648;
Best Local Similarity 86.0%; Pred. No. 2e-100;
Matches 555; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 2585 AATCATCTTAATTAGACTGACTGACAAATAGGGCTACAGAGTACCCTGAAAATTGAGTT 2644
DB   3  AATCTGTCTTAATTGAGATCTGCACAGATAGCAACGTAGAAAGTAAACAGACAATCTTAGAGAT 62

QY 2645 TCTCATTTCCCTCCATTAATACAGTGACGTGGAGTAAAGATTCAATCATCTGCAATTTGGAG 2704
DB   63  TCTCTCTTTCCCTCGATTAACGGTCAATTTGGCAATGAAGGTTCAATCACACCAANTTGGAG 122

QY 2705 ATGTTCTTGGCTCTTCATCAATAGGCTTAGGCTCATTTGATTCGGATGCCATTAGGCTGTG 2764
DB  123 ATATCTTTGGTGTCTTCATCAATAGGCTTATTTTCACTGATCCGATGCTTATGATATGTG 182

QY 2765 GTGAACAGAACATGATAAATTTTGCCTCCAATATTTTACATTTTGGATTAATCTGACTAAAA 2822
DB  183 GTGAACAGAACATGATTAATTTTGTCTCCAATATTTTACATTTTGGATTAATCTGACTAAAC 242

QY 2825 AGAAACAACTGACAGATTAATTGAAAGAAAAGCTCTTCAATTATGAGGAAAGTTTAC 2888
DB  243 AGAAACAGCTGACAGTAAATTTAAAGAAAAGSCCTTTCAATATATGAGGCAAAGTTTACC 302

QY 2885 AGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGTCTTTTGGCAATTATGACC 2944
DB  303 AAAGGAGCTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGTCTTTTGGCAACTATGACT 362

QY 2945 CTTCTGGAGACATTTGGTGTGACCTTTTGTTTTAAAGTGTTCCTTGAAGCCGATCTTT 3004
DB  363 CTTCTGGAGACATTTGGTGTGACCTTTTGTTTTAAAGTGTTCCTTGAAGCCGATTAATCT 422

QY 3005 ACATAGATATTTGATCAGAAATGTGTTTACACAGAAATATACATCTTGGCTTAAAGACATACGA 3066
DB  423 ATATAGATATTTGATCAGAAATGTGTTTACACAGAAATATATCTTGGCTTAAATGACATATAGA 482

QY 3065 AATCCAACGGGATTTTGGGATCCAGAAAGAGTATTCATAGAGAGCTTCAAGGTGACCA 3124
DB  483 AATCCAATGGGATTTTGGGATCCAGAAAGAGTATTCATAGAGAGCTTCAAGGTGACCA 542

QY 3125 ATAAGAATCCAGTAAACCTTACAGCTTAATATTGACTTCTCTCTGGGATATAGAAAGT 3184
DB  543 CCABAAGCCAGTAAACCTTACAGCTTAATATTGACTTCTCTCTGGGATATAGAAAGT 602

QY 3185 ATCAGCTTAACATTTGATGTGCAGAGTGTATTCATTTTGGAGT 3229
DB  603 ATCAGCTTAATATGATGTATCAGAACTCAATCAAGTTTGGAGAT 647

```



CB158337  
LOCUS CB158337 613 bp mRNA linear EST 29-JAN-2003  
DEFINITION K-EST0217654 L18POOL1n1 Homo sapiens cDNA clone L18POOL1n1-10-H06  
ACCESSION CB158337  
VERSION CB158337.1 GI:28143473  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 613)  
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
Kim,Y.S.  
21C Frontier Korean EST Project 2001  
JOURNAL Unpublished  
COMMENT Contact: Kim YS  
Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology  
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongsung@mail.kribb.re.kr  
Plate: 10 row: H column: 06  
High quality sequence stop: 613.  
Location/Qualifiers  
1..613  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="L18POOL1n1-10-H06"  
/cell\_line="SNU-354+Cho-CK+Choi-CK+HLK-3"  
/lab\_host="Top10F"  
/clone\_lib="L18POOL1n1"  
/note="Organ: Liver; Vector: pTTT3-Pac; Site\_1: EcoRI;  
Site\_2: NotI; The library was contributed by the Soares  
Laboratory and it was constructed as described by Bonaldo,  
M.F., Lennon, G. and Soares, M.B. (1996), Genome Research  
6(9): 791-806. RNA was prepared from harvested cell  
culture."  
BASE COUNT 183 a 111 c 142 g 177 t  
ORIGIN  
Query Match 11.6%; Score 498; DB 14; Length 613;  
Best Local Similarity 99.8%; Pred. No. 1.2e-99;  
Matches 509; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
QY 3652 CTTGCTGTGCTACAGCCCAATGACGATTATTTCCGCAAAATGTTGGATTGCTATT 3711  
DB 104 CTTGCTGTGCTACAGCCCAATGACGATTATTTCCGCAAAATGTTGGATTGCTATT 163  
QY 3712 TGTGAGCTCAATGTTGTATATTAATGTAAGAGCTTCTGGGCTCTTCTGAAAGACGAAGATCT 3771  
DB 164 TGTGAGCTCAATGTTGTATATTAATGTAAGAGCTTCTGGGCTCTTCTGAAAGACGAAGATCT 223  
QY 3772 ATCCAAATCAAGAAAGCCTTTGATTAGATGTTGCTGTAAAGAAATAAAGATGATCTC 3831  
DB 224 ATCCAAATCAAGAAAGCCTTTGATTAGATGTTGCTGTAAAGAAATAAAGATGATCTC 283  
QY 3832 AATCATGTGATTTGATGATGTGTACAAAGCTTTCCGGGCCCGGGTGTAGAGTGGCATGGCT 3891  
DB 284 AATCATGTGATTTGATGATGTGTGTACAAAGCTTTCCGGGCCCGGGTGTAGAGTGGCATGGCT 343  
QY 3892 CTTATGAAAGTTAAGCTATTTAAGTGTCTTTATGTCCTTTAGAACCAATTTCTGTAGC 3951  
DB 344 CTTATGAAAGTTAAGCTATTTAAGTGTCTTTATGTCCTTTAGAACCAATTTCTGTAGC 403  
QY 3952 GAGACAGTGAAGAAAGTGAATGTGATCAATGAAACCTCAATTTAAGATTCTGTA 4011  
DB 404 GAGACAGTGAAGAAAGTGAATGTGATCAATGAAACCTCAATTTAAGATTCTGTA 463  
QY 4012 AATGAACCCAGTTTGTGTATATTTCTGCTGTGAGAAACCTTTAAAGTTCAATATCC 4071

DB 464 AATGAACCCAGTTTGTGTATATTTCTGCTGTGAGAAACCTTTAAAGTTCAATATCC 523  
QY 4072 CAAG-ATGCTTACAGTCCATATGATGATTTACTATAGCAAGAGACGCGGTAGAG 4130  
DB 524 CAAGATGCTTACAGTCCATATGATGATTTACTATAGCAAGAGACGCGGTAGAG 583  
QY 4131 TTACACTCTGAGTGAAGTGTCTCTCTG 4160  
DB 584 TTACACTCTGAGTGAAGTGTCTCTCTG 613  
RESULT 14  
BX472602 515 bp mRNA linear EST 12-JUN-2003  
LOCUS BX472602  
DEFINITION DKFZp686C02145\_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone  
ACCESSION DKFZp686C02145.5, mRNA sequence.  
VERSION BX472602  
KEYWORDS BX472602.1 GI:31666864  
SOURCE EST.  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 515)  
Bloecker,H., Boecker,M., Mewes,H.W., Weill,B., Amid,C., Osanger,A.,  
Fobbo,G., Han,M. and Wiemann,S.  
EST (Bloecker,H., Boecker,M., Mewes,H.W., Weill,B., Amid,C., et al.)  
JOURNAL Unpublished  
COMMENT Contact: Bloecker H  
MIPS  
Ingolstaedter Landstr.1, D-85764 Neuberg, Germany  
This is the 5' sequence of the clone insert  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
Braunschweig/Germany) within the cDNA sequencing consortium of the  
German Genome Project.  
a1 sequence also available.  
This clone (DKFZp686C02145) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.  
Location/Qualifiers  
1..515  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DKFZp686C02145"  
/tissue="human skeletal muscle"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_lib="686 (synonym: hlcc3)"  
/note="Vector: pTriplex2; Site\_1: SfiI; Site\_2: SfiI;  
cDNA-collection"  
BASE COUNT 130 a 142 c 120 g 123 t  
ORIGIN  
Query Match 11.6%; Score 496; DB 13; Length 515;  
Best Local Similarity 100.0%; Pred. No. 3.3e-99;  
Matches 496; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AATGAGGGGCCACGCTCTGACCGCGGCCACCTCTGCTGTGACGCGCGGCTG 60  
DB 20 AATGAGGGGCCACGCTCTGACCGCGGCCACCTCTGCTGTGACGCGCGGCTG 79  
QY 61 GCGTGGCTCCCGGCGCTCGGTTTCTGTGTGACAGCCCGAGGATCATGAGCCCGAGGA 120  
DB 80 GCGTGGCTCCCGGCGCTCGGTTTCTGTGTGACAGCCCGAGGATCATGAGCCCGAGGA 139  
QY 121 AATGTACTATTGGGCTGTGAGCTTGTGAACATGCTCTTCAACAGTATCTGTGAAGCG 180  
DB 140 AATGTACTATTGGGCTGTGAGCTTGTGAACATGCTCTTCAACAGTATCTGTGAAGCG 199

QY 181 GAGTCTCAAGACATCAACCTCACTGCTCTGCTCTGGAAGAGAGAGCTTT 240  
 DB 200 GAGCTGCTCAAGACATCAACCTCACTGCTCTGCTCTGGAAGAGAGAGCTTT 259  
 QY 241 GAAAGAGCTCTTTAAGACACTTCTCTGCTCTGCTCTGGAAGAGAGAGCTTT 300  
 DB 260 GAAAGAGCTCTTTAAGACACTTCTCTGCTCTGCTCTGGAAGAGAGAGCTTT 319  
 QY 301 ATTATGAGCTACGCTGTAACCGGACGTACCGAGATGAGATTTATTTATGATACC 360  
 DB 320 ATTATGAGCTACGCTGTAACCGGACGTACCGAGATGAGATTTATTTATGATACC 379  
 QY 361 CGCTTATGCTTGAAGACCAAGATATCTGCTTCAATCAACAGACAGGCTTATAC 420  
 DB 380 CGCTTATGCTTGAAGACCAAGATATCTGCTTCAATCAACAGACAGGCTTATAC 439  
 QY 421 AAGCAAGCAAGAGAGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
 DB 440 AAGCAAGCAAGAGAGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 499  
 QY 481 AAAACCTCTTAAACA 496  
 DB 500 AAAACCTCTTAAACA 515

## RESULT 15

LOCUS BX370169 919 bp mRNA linear EST 08-MAY-2003  
 DEFINITION BX370169 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
 clone CS0D1081YB18 5-PRIME, mRNA sequence.

ACCESSION BX370169 GI:30455906  
 VERSION BX370169.1  
 KEYWORDS EST.

SOURCE Homo sapiens (human)

## ORGANISM

Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 919)  
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished  
 COMMENT Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 Evry cedex - France  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 Library was constructed by Life Technologies, a division of  
 Invitrogen. This sequence belongs to sequence cluster 3578.f. For  
 more information about this cluster, see  
 http://www.genoscope.cns.fr/  
 cgi-bin/cluster.cgi?seq=CS0BAG038ZH04\_CS03584\_1&cluster=3578.f.  
 Contact : Feng Liang Email : fliang@lifetech.com URL :  
 http://fulllength.invitrogen.com/Invitrogen Corporation 1600  
 Paradey Avenue Genoscope sequence ID : CS0BAG038ZH04\_CS03584\_1.

## FEATURES

source

1. 919  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0D1081YB18"  
 /tissue\_type="PLACENTA COT 25-NORMALIZED"  
 /clone\_id="Homo sapiens PLACENTA COT 25-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-oligo(dT)  
 primer. Five prime end enriched, double-strand cDNA was  
 digested with Not I and cloned into the Not I and EcoR V  
 sites of the pCMVSPORT 6 vector. Library was normalized."  
 BASE COUNT 285 a 213 c 177 g 242 t  
 ORIGIN

Query Match 11.5%; Score 493.8; DB 13; Length 919;  
 Best Local Similarity 94.4%; Pred. No. 1.1e-98;  
 Matches 577; Conservative 0; Mismatches 27; Indels 7; Gaps 6;  
 QY 722 TGAATCTAAGCATTTAATGTCATCAAGCAAGATATATATGGAAGCCAGTGA 781

DB 4 TGAATCTAAGCATTTAATGTCATCAAGCAAGATATATATGGAAGCCAGTGA 63  
 QY 782 AAGAGACGTAAAGCTTACATTTTATCTTATCTTTGGGAGAGAGAGATTTA 841  
 DB 64 AAGAGACGTAAAGCTTACATTTTATCTTATCTTTGGGAGAGAGAGATTTA 123  
 QY 842 CAAAGACATTTAAGATTAATGATGATGCAAACTCTCTTTTAAATGATGAAGATGAAA 901  
 DB 124 CAAAGACATTTAAGATTAATGATGATGCAAACTCTCTTTTAAATGATGAAGATGAAA 183  
 QY 902 ATGTAATGATTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
 DB 184 ATGTAATGATTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 243  
 QY 961 GTAGAAATTTTAAACCAAGTACAGATGATGATGATGATGATGATGATGATGATGATGAT 1020  
 DB 244 GTAGAAATTTTAAACCAAGTACAGATGATGATGATGATGATGATGATGATGATGATGAT 303  
 QY 1021 AATGCTTCTTCAAGCAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1079  
 DB 304 AATGCTTCTTCAAGCAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 363  
 QY 1080 GAAGCCATCTCTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCA 1138  
 DB 364 GAAGCCATCTCTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCA 423  
 QY 1139 TGAATCTTGAAGAGAGAGATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1198  
 DB 424 TGAATCTTGAAGAGAGAGATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 483  
 QY 1199 AGTACTGAGACGGATCTTAAAGTGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1258  
 DB 484 AGTACTGAGACGGATCTTAAAGTGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 543  
 QY 1259 A-TACTGTCCCCCAAGTGAATTT-TTAAAGTTGAATTT-CCCATCTCTGAGAGATTC 1314  
 DB 544 ATTACTGTCCCCCAAGTGAATTT-TTAAAGTTGAATTT-CCCATCTCTGAGAGATTC 603  
 QY 1315 AGTGAAGTACA 1325  
 DB 604 AGCGAGCTTCA 614

Search completed: January 16, 2004, 20:57:23  
 Job time : 7735 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2004, 12:38:20 ; Search time 14986 Seconds

(without alignments)  
11702.894 Million cell updates/sec

Title: US-10-020-095-3

Perfect score: 4287

Sequence: 1 atgcagggccacgcgcctcct.....ttatggaacttgcgtcgtga 4287

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

GenEmbl:\*

1: gb\_ba:\*

2: gb\_hcg:\*

3: gb\_in:\*

4: gb\_om:\*

5: gb\_ov:\*

6: gb\_pac:\*

7: gb\_ph:\*

8: gb\_pl:\*

9: gb\_pr:\*

10: gb\_ro:\*

11: gb\_ses:\*

12: gb\_sy:\*

13: gb\_un:\*

14: gb\_vi:\*

15: em\_ba:\*

16: em\_fun:\*

17: em\_hum:\*

18: em\_in:\*

19: em\_mu:\*

20: em\_om:\*

21: em\_or:\*

22: em\_ov:\*

23: em\_pac:\*

24: em\_ph:\*

25: em\_pl:\*

26: em\_ro:\*

27: em\_ses:\*

28: em\_un:\*

29: em\_vi:\*

30: em\_hcg\_hum:\*

31: em\_hcg\_hiv:\*

32: em\_hcg\_ocher:\*

33: em\_hcg\_mus:\*

34: em\_hcg\_pin:\*

35: em\_hcg\_rod:\*

36: em\_hcg\_mam:\*

37: em\_hcg\_vrt:\*

38: em\_sy:\*

39: em\_higo\_hum:\*

40: em\_higo\_mus:\*

41: em\_higo\_other:\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4224.4	98.5	4761	6	AX534934
2	4224.4	98.5	5883	6	AF410459 Homo sapi
3	4224.4	98.5	5895	6	AX534938
4	4222.8	98.5	4761	6	AX534936
5	4222.8	98.5	5895	6	AX534940
6	4221.4	98.5	4335	6	AX537612
7	4219.8	98.4	4335	6	AX537612
8	3187.4	74.4	3535	6	AX534942
9	3185.8	74.3	3535	6	AX534944
10	2709.2	63.2	4735	10	AY083458
11	2709.2	63.2	5644	10	BC052443
12	2448.2	57.1	2938	9	HS080581
13	1969.4	45.9	2273	9	AK095888
14	730.8	17.0	1300	6	AX083146
15	232	5.4	163577	9	AL590428
16	230	5.4	176144	2	AC012408
17	219.6	5.1	5603	3	CIN431688
18	203.8	4.8	4678	3	D83196
19	182	4.2	251250	2	AC097023
20	177.4	4.1	91419	9	AL591480
21	175	4.1	149597	2	AC034271
22	174	4.1	4527	6	AX054953
23	174	4.1	4560	6	AX054954
24	171.6	4.0	2629	3	AY118302
25	167.6	3.9	4715	5	GGOVOM
26	167	3.9	4578	3	AY122084
27	167	3.9	4579	3	DME269539
28	160	3.7	144589	2	AC112668
29	160	3.7	191258	2	AC118254
30	155.4	3.6	4620	10	RNA113
31	154.8	3.6	176144	2	AC012408
32	154.2	3.6	4636	10	RAT113A
33	152.2	3.6	4220	3	DME269538
34	151.2	3.5	4575	10	GPIMSPB
35	149.8	3.5	2501	10	RAT113
36	146.4	3.4	4950	5	XELENDO
37	143.8	3.4	122013	3	AC092395
38	143.8	3.4	164414	3	AC092395
39	143.8	3.4	168503	3	AC115484
40	143.8	3.4	275699	3	AE003649
41	143.8	3.4	337178	3	DROSADH08
42	140.8	3.3	191258	2	AC118254
43	136	3.2	4623	10	BC036299
44	136	3.2	4672	10	BC051037
45	134.8	3.1	2689	5	AB026130

#### ALIGNMENTS

RESULT 1	AX534934	4761 bp	DNA	linear	PAT 22-NOV-2002
LOCUS	AX534934				
DEFINITION	Sequence 1 from Patent WO02070696.				
ACCESSION	AX534934				
VERSION	AX534934.1	GI:25261474			
KEYWORDS					
SOURCE					
ORGANISM	Homo sapiens (human)				
	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE					
AUTHORS	Schuh, A. and Sutherland, R. D.				
TITLE	Cd109 nucleic acid molecules polypeptides and methods of use				
JOURNAL	Patent: WO 02070696-A 1 12-SEP-2002;				

Pred. No. is the number of results predicted by chance to have a

FEATURES  
source Schuh, Andre (CA) ; Sutherland, Robert D. (CA)

## CDS

1..4761  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
113..4450  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD57231.1"  
/db\_xref="gi:25261475"

/translation="MQGPPPLTAHLLCVCTAALAVAPRPFVLTAGIIRPGNVTI  
GVELLEHCPQVTVKAEKLLKTSNLTVSLAEVFEKSGFKTLTSLPLNSADELY  
ELRYVGRDEILFNSNTRLSFEKRIISVFIQDRAKYKPKQKRFIVLIFSDPKFY  
KTSINILIKDEKSNLLOMISQOSDLGIVISKIPLSDMSIQVNDQYVYOS  
FOSEYVLPKFEVTLQPLCYGSKNKLNGITAKYTKYGVKVDLTLTFESFMK  
KKNITKTUKINGSANFSPNDEMNKNSNGLSSEYDLSPRGVEILITTYESVTEI  
SRNVSTNVFKQHDITIEFDYITVLKSLNFTAVKTRADGNQLTLEERKNVIT  
VTORNYTEWGSNSNQKMEAVOKINVTPOSGTFKLEPILDESELQKAVFLGS  
KSMASVHSLFSPKTYIQLKTRDENIKVSGPFLVVGKRLKELSYMVNSQOLVA  
VGNKSTWFLSLPNSMTPKACIVYVIEDGELISDYLKIPVLVFNKIKJWMSYV  
KAPSEKSLRISVTOPDSIVGIVAVKSVLMAASNDITMENVTHELKLYNGYVIG  
MPNMSFAPFQEGMLVLTDAUTDYIDGVYDNAEYARFMEENEGHIVIDHDSLS  
SPHVRKHPETWILMDTWGRIYIOEFVTVPDSITSMVATGVISDLGIGLITTV  
ELQAFQPFPIFLNLPYSVIRGEBALEITIFNYLKDAIEVYIIEKSDPIDLMTSE  
INATGHOQTLVSPEDGATVLPPIRPTLGEIPITVVALSPADAVTOMILVABEII  
EKXSOSILDLTNRLQSTLKTLSFSPNVTGSEVOITAGVIGPSINGLASL  
IRMPYCGEONINPAPNIIYLDYLTKKKOLLNLKKAALSFMGQYORLELYORNG  
SPSAPNVDPSGSTMSPFVLRCELEADPIDIDQNTLHRTYTLKQHSNENFMP  
GRVHSELQGNKSPVLTAVITVSLGRIKQPNIDQESIHLESEFSGISDNT  
LALITVALSVSGSPAKALMLTWRAEGEQMOWSSEKLSDSQPSRLDEVA  
YALSHLPQFQSEGIPIRMWLSRQNSLQSGASTODITVALKALSFALMETERN  
IOVTVGSSPSPVKFLIDHNLRLLOTAELAVQPAVINSANGFQALCOLNVN  
VKAGSSRRRSRSIONEAFDLDAVKENKDLNHDINVCFSFSGPGRGMALEVN  
LSGMVPSBEASISSETYKUYEDGKLANLYIDSYNEFPCNIPAYNPKYSNTQDS  
VSIYDYEPRQAVRSINSEVSKSSCDLVQGCRCPEBDASGSHHSVIFLFCFK  
LLYFMEIWL"

BASE COUNT 1420 a 912 c 1016 g 1413 t  
ORIGIN

Query Match 98.5%; Score 4224.4; DB 6; Length 4761;  
Best Local Similarity 98.8%; Pred. No. 0; 4450  
Matches 4286; Conservative 0; Mismatches 1; Indels 51; Gaps 1;

1 ATGCAAGGCGCCACGCTCTGACCGCGCCACGCTCTGAGTGTGACCGCGCGCTG 60  
113 ATGCAAGGCGCCACGCTCTGACCGCGCCACGCTCTGAGTGTGACCGCGCGCTG 172  
61 GCCGTGCTCCCGGCTCTGCTGTTCTGTGACAGCCCGAGGATCATCAGGCCGGA 120  
173 GCCGTGCTCCCGGCTCTGCTGTTCTGTGACAGCCCGAGGATCATCAGGCCGGA 232  
121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCAACAGGTGACTGTAAAGCG 180  
233 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCAACAGGTGACTGTAAAGCG 292  
181 GAGCTGCTCAAGACAGCATCAACCTCACTGCTCTGCTCTGGAAGAGAGAGAGTCTTT 240  
293 GAGCTGCTCAAGACAGCATCAACCTCACTGCTCTGCTCTGGAAGAGAGAGAGTCTTT 352  
241 GAAAAAGGCTTTTAAAGACACTTACTTTCATCACTACTCTGAAACAGTGCAGATGAG 300  
353 GAAAAAGGCTTTTAAAGACACTTACTTTCATCACTACTCTGAAACAGTGCAGATGAG 412  
301 ATTATAGAGTACGTGTAAACCGAGAGTACCCAGAGATGAGATTTTATTTCTTAATAGTACC 360  
413 ATTATAGAGTACGTGTAAACCGAGAGTACCCAGAGATGAGATTTTATTTCTTAATAGTACC 472  
361 CGCTTATCATTTGAGACCAAGAGAAATCTGTCTTCAATTCAGAACAGAGGCTTTATAC 420  
473 CGCTTATCATTTGAGACCAAGAGAAATCTGTCTTCAATTCAGAACAGAGGCTTTATAC 532  
421 AAGCCAAAGCAAGAGTGAAGTTTCGCAATGTGTACACTCTTCTCAGATTTTAAAGCTTTAC 480

533 AAGCCAAAGCAAGAGTGAAGTTTCGCAATGTGTACACTCTTCTCAGATTTTAAAGCTTTAC 592  
481 AAAACGCTTTAAACATCTCTCATTTAAGAGACCCCAATCAAAATTTGATCCAAAGTGGTTG 540  
593 AAAACGCTTTAAACATCTCTCATTTAAGAGACCCCAATCAAAATTTGATCCAAAGTGGTTG 652  
541 TCACAAACAAAGTATCTTGGAGTCAATTTCCAAACATTTTCAGCTATCTTCCATCCAAATA 600  
653 TCACAAACAAAGTATCTTGGAGTCAATTTCCAAACATTTTCAGCTATCTTCCATCCAAATA 712  
601 CTGCTGACCTGCTCTATTCAGATTCAGAGTAATGACACGACATATTAATCATTTTCAG 660  
713 CTGCTGACCTGCTCTATTCAGATTCAGAGTAATGACACGACATATTAATCATTTTCAG 722  
661 GTTTCAGAAATATGATATTAACCAAAATTTGAAAGTACCTTGGACACACATATATGTTCT 720  
773 GTTTCAGAAATATGATATTAACCAAAATTTGAAAGTACCTTGGACACACATATATGTTCT 832  
721 ATGAATTTCTAAGCATTTTAAATGATGATACATCAACGCAAGATATCATATGGAAGCAGTG 780  
833 ATGAATTTCTAAGCATTTTAAATGATGATACATCAACGCAAGATATCATATGGAAGCAGTG 892  
781 AAAGGACGTAACGCTTACATTTTACCTTTATCCTTTTGGGAAAGAAATATAT 840  
893 AAAGGACGTAACGCTTACATTTTACCTTTATCCTTTTGGGAAAGAAATATAT 952  
841 ACAAACAAATTTAAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
953 ACAAACAAATTTAAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1012  
901 AATGTAATGATTTCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
1013 AATGTAATGATTTCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1072  
961 GTGAAATTTTAAACCAAGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020  
1073 GTGAAATTTTAAACCAAGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1132  
1021 AATGTCCTTCAAGCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080  
1133 AATGTCCTTCAAGCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1192  
1081 AAGCCATCTCTCAACTTCAAGCAACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140  
1193 AAGCCATCTCTCAACTTCAAGCAACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1252  
1141 ACTCTTGAAGAAAGAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200  
1253 ACTCTTGAAGAAAGAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1312  
1201 TACTGAGAGGAGTCTTAACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1260  
1313 TACTGAGAGGAGTCTTAACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1372  
1261 ACTGTCCTCCCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1320  
1373 ACTGTCCTCCCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1432  
1321 CTACAGTTGAAGGCTTATTTCTTGGTATGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1380  
1433 CTACAGTTGAAGGCTTATTTCTTGGTATGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1492  
1381 AAGTCTCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1440  
1493 AAGTCTCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1552  
1441 TCGCCTTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1500  
1553 TCGCCTTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1612  
1501 GTATCAAGGAGACAGTTGGTGGCTGTAGAGAAACAAATTCACAAATGTTCTCTTTTAA 1560

Db 1613 GTATCAGGGGACAGTTGGCTGTAGAAAACAAATTCACAAATGCTCTTTACA 1672  
Qy 1561 CCAGAAAATTTCTTGGACTCCAAAAGCTGTGTAATTTGTATTAATTAATGAATGATGG 1620  
Db 1673 CCAGAAAATTTCTTGGACTCCAAAAGCTGTGTAATTTGTATTAATTAATGAATGATGG 1732  
Qy 1621 GAAATTTATAGTATGTTCTTAAAAATCTCTGTTCAGCTGTTTAAAAATTAAGTAAG 1680  
Db 1733 GAAATTTATAGTATGTTCTTAAAAATCTCTGTTCAGCTGTTTAAAAATTAAGTAAG 1792  
Qy 1681 CTATATTTGAGTAAGTGAAGCTGAACCATCTGAGAAAGTCTCTTTAGATCTCTGG 1740  
Db 1793 CTATATTTGAGTAAGTGAAGCTGAACCATCTGAGAAAGTCTCTTTAGATCTCTGG 1852  
Qy 1741 ACACAGCTGATCTCATAGTGGATTTGAGCTTTGACAAAAGTGTAAATCTGATGAAT 1800  
Db 1853 ACACAGCTGATCTCATAGTGGATTTGAGCTTTGACAAAAGTGTAAATCTGATGAAT 1912  
Qy 1801 GCTCTTAATGATATTACAAATGAAAAATGTGTCATGAGTTGAACTTTATAACACAGGA 1860  
Db 1913 GCTCTTAATGATATTACAAATGAAAAATGTGTCATGAGTTGAACTTTATAACACAGGA 1972  
Qy 1861 TATTTATTTAGGATGTTCAATGAATTTCTTTGACAGTCTTTCAGAAATGTGACTGGGTA 1920  
Db 1973 TATTTATTTAGGATGTTCAATGAATTTCTTTGACAGTCTTTCAGAAATGTGACTGGGTA 2032  
Qy 1921 TTGACAGATGCAAACTCCAGAAAGATTATTTGATGTTGTTATGACAAATGACAAATAT 1980  
Db 2033 TTGACAGATGCAAACTCCAGAAAGATTATTTGATGTTGTTATGACAAATGACAAATAT 2092  
Qy 1981 GCTGAGAGTTTATGAGAGAAAAATGAGACATATTTGATATTTCAATGCTTTCTTTG 2040  
Db 2093 GCTGAGAGTTTATGAGAGAAAAATGAGACATATTTGATATTTCAATGCTTTCTTTG 2152  
Qy 2041 GGTAGACAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACAC 2100  
Db 2153 GGTAGACAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACAC 2212  
Qy 2101 ATGGGTTACAGAGATTACCAAGAAATTTGAAGTAACTGTACCTGATCTGTACCTCTGG 2160  
Db 2213 ATGGGTTACAGAGATTACCAAGAAATTTGAAGTAACTGTACCTGATCTGTACCTCTGG 2272  
Qy 2161 GTGGCTACTGTTTTGTGATCTCTGAGACCTGGGCTTTGACCTAACACTACTCCAGTG 2220  
Db 2273 GTGGCTACTGTTTTGTGATCTCTGAGACCTGGGCTTTGACCTAACACTACTCCAGTG 2332  
Qy 2221 GAGCTCCAAAGCTTCCAACTATTTTCAATTTTGTGAATCTTCCCTACTCTGTATCAGA 2280  
Db 2333 GAGCTCCAAAGCTTCCAACTATTTTCAATTTTGTGAATCTTCCCTACTCTGTATCAGA 2392  
Qy 2281 GGTGAAGAAATTTGCTTTGGAATACTAATTAATTAATTTGAAAGATGCCCTGAGGT 2340  
Db 2393 GGTGAAGAAATTTGCTTTGGAATACTAATTAATTAATTTGAAAGATGCCCTGAGGT 2452  
Qy 2341 AAGGTAAATCAATGAGAAAAAGTGAACAAATTTGATTAATCTAATGACTTCAAGTAAATTAAT 2400  
Db 2453 AAGGTAAATCAATGAGAAAAAGTGAACAAATTTGATTAATCTAATGACTTCAAGTAAATTAAT 2512  
Qy 2401 GGCACAGGCGACACAGACAGCCCTTGTGTTCCAGTGAAGATGGGCAACTGTCTTTT 2460  
Db 2513 GGCACAGGCGACACAGACAGCCCTTGTGTTCCAGTGAAGATGGGCAACTGTCTTTT 2572  
Qy 2461 GCCATCAGGCCCAACATCTGGAGAAAAATTTCTATTCACAGTCAAGCTTTTCAACCACT 2520  
Db 2573 GCCATCAGGCCCAACATCTGGAGAAAAATTTCTATTCACAGTCAAGCTTTTCAACCACT 2632  
Qy 2521 GCTTCTGATGCTGTCAACCCAGATGATTTTAAAGCTGAAGAAATGAAGAAATCAATAT 2580  
Db 2633 GCTTCTGATGCTGTCAACCCAGATGATTTTAAAGCTGAAGAAATGAAGAAATCAATAT 2692  
Qy 2581 TCACAAATCATCTTATTTAGACTTGACAAATAGGCTACAGAGTACCTGAAAACTTTG 2640  
Db 2693 TCACAAATCATCTTATTTAGACTTGACAAATAGGCTACAGAGTACCTGAAAACTTTG 2752

Qy 2641 AGTTTCTGATTTCTCTTAATACAGTGAAGTGGAGTGAAGAGTTCAGATCACTGCAATT 2700  
Db 2753 AGTTTCTGATTTCTCTTAATACAGTGAAGTGGAGTGAAGAGTTCAGATCACTGCAATT 2812  
Qy 2701 GGAAGATGTTCTTGGTCTTCCATCAATGAGCTTACCTCAATGATTTGGATGCTTATGGC 2760  
Db 2813 GGAAGATGTTCTTGGTCTTCCATCAATGAGCTTACCTCAATGATTTGGATGCTTATGGC 2872  
Qy 2761 TGTGTGAACGAACATGATTAATTTGCTCCAAATATTTCAATTTGATTAATGATCT 2820  
Db 2873 TGTGTGAACGAACATGATTAATTTGCTCCAAATATTTCAATTTGATTAATGATCT 2932  
Qy 2821 AAAAAAACAACCTGACAGATTAATTTGAAGAAAAAGCTCTTCAATTAATGAGCAAGT 2880  
Db 2933 AAAAAAACAACCTGACAGATTAATTTGAAGAAAAAGCTCTTCAATTAATGAGCAAGT 2992  
Qy 2881 TACAGAGAGAACTTCTATCAGAGGGAATGCTCTTCACTGCTTTGGGAATTAAT 2940  
Db 2993 TACAGAGAGAACTTCTATCAGAGGGAATGCTCTTCACTGCTTTGGGAATTAAT 3052  
Qy 2941 GACCTTCTGGAGACATTTGGTGTCAAGCTTTTGTAAAGTGTCTTGAAGCCGAT 3000  
Db 3053 GACCTTCTGGAGACATTTGGTGTCAAGCTTTTGTAAAGTGTCTTGAAGCCGAT 3112  
Qy 3001 CCTTACATAGATTAATGATCAGAAATGTTTACACAGAACTACCTTGCTTAAAGACAT 3060  
Db 3113 CCTTACATAGATTAATGATCAGAAATGTTTACACAGAACTACCTTGCTTAAAGACAT 3172  
Qy 3061 CAGAAATCCAAACGATGAATTTTGGATCCAGAAAGATGATTCATAGTGAAGCTTCAAGT 3120  
Db 3173 CAGAAATCCAAACGATGAATTTTGGATCCAGAAAGATGATTCATAGTGAAGCTTCAAGT 3232  
Qy 3121 GGCAATTAAGTCCAGTAACTTACACCTTAATTTGTAATCTTCTCTGGGAATTAAGA 3180  
Db 3233 GGCAATTAAGTCCAGTAACTTACACCTTAATTTGTAATCTTCTCTGGGAATTAAGA 3292  
Qy 3181 AAGTATCAAGCTTAACATTTGATGAGAGAGTCAATTCATTTTGGAGTCTGAATTCAGT 3240  
Db 3293 AAGTATCAAGCTTAACATTTGATGAGAGAGTCAATTCATTTTGGAGTCTGAATTCAGT 3352  
Qy 3241 AGAGGAATTTCAACAATTAATACCTTACCCCTTAATTAATGCAATTTGATCACTGAGG 3300  
Db 3353 AGAGGAATTTCAACAATTAATACCTTACCCCTTAATTAATGCAATTTGATCACTGAGG 3412  
Qy 3301 AGTCTTAAGGAGAGAGAGCTTTGAATGCTGACTTGAAGAGCAAGAAAGAGTGGC 3360  
Db 3413 AGTCTTAAGGAGAGAGAGCTTTGAATGCTGACTTGAAGAGCAAGAAAGAGTGGC 3472  
Qy 3361 ATGCAATTTCTGGGTCTATCAGAGTCCAACTTTCTGACTCTGGACACAGCTCCCTG 3420  
Db 3473 ATGCAATTTCTGGGTCTATCAGAGTCCAACTTTCTGACTCTGGACACAGCTCCCTG 3532  
Qy 3421 GATATTTGAAGTTGACGCTTATGACTGCTCTCACACTTTCACAATTTCAAGCTTCTGAG 3480  
Db 3533 GATATTTGAAGTTGACGCTTATGACTGCTCTCACACTTTCACAATTTCAAGCTTCTGAG 3592  
Qy 3481 GGAATCCCAATTAATGAGTGGCTAAAGAGGCAAGAAATTAAGTGGTGGTCTTTCATCT 3540  
Db 3593 GGAATCCCAATTAATGAGTGGCTAAAGAGGCAAGAAATTAAGTGGTGGTCTTTCATCT 3652  
Qy 3541 ACTCAGATTAACCTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAAGCCCTAATGAATCA 3600  
Db 3653 ACTCAGATTAACCTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAAGCCCTAATGAATCA 3712  
Qy 3601 GAAAGCAAAATATCAAGTACCGTGAAGGGGCTTACTGACCAAGTCTCTCTCTCTCTCT 3651  
Db 3713 GAAAGCAAAATATCAAGTACCGTGAAGGGGCTTACTGACCAAGTCTCTCTCTCTCTCT 3772  
Qy 3652 -----CTTGTGTGATCAGGCA 3669  
Db 3773 CTGATTTGACACACAAACCGTTACTCTTCAAGACAGAGAGTGTGTGTACAGGCA 3832

```

Qy 3670 ATGCGAGTAAATATTTTCGCAAAATGTTTGGATTTGCTATTTTGCAGCTCAATGTTGTA 3729
Db 3633 ATGCGAGTAAATATTTTCGCAAAATGTTTGGATTTGCTATTTTGCAGCTCAATGTTGTA 3892
Qy 3730 TATATATGTAAGGCTTTCTGGGCTTTCTAGAAAGAGAGATTTATCCAAATCAAGAGCC 3789
Db 3693 TATATATGTAAGGCTTTCTGGGCTTTCTAGAAAGAGAGATTTATCCAAATCAAGAGCC 3952
Qy 3790 TTTGATTTAGATGTTGCTGTAAAGAAAATAAGATGATCTCAATCAATGTTGAT 3849
Db 3953 TTTGATTTAGATGTTGCTGTAAAGAAAATAAGATGATCTCAATCAATGTTGAT 4012
Qy 3850 GTGTGTACAGCTTTTCGGGCGCGGTAGAGTGCATGCTCTTATAGAAATTAACCTTA 3909
Db 4013 GTGTGTACAGCTTTTCGGGCGCGGTAGAGTGCATGCTCTTATAGAAATTAACCTTA 4072
Qy 3910 TTAAGTGGCTTTATGAGTCCCTTCAGAGCAATTTCTGACCGAGACAGTGAAGAGTG 3969
Db 4073 TTAAGTGGCTTTATGAGTCCCTTCAGAGCAATTTCTGACCGAGACAGTGAAGAGTG 4132
Qy 3970 GAATATGATCATGGAAGAACTCAACCTCTATTAGATCTGTAAATGAAACCAAGTTTGT 4029
Db 4133 GAATATGATCATGGAAGAACTCAACCTCTATTAGATCTGTAAATGAAACCAAGTTTGT 4192
Qy 4030 GTTATATTTCTGCTGTGAGAACTTTAAAGTTTCAATACCAAGATGCTTCAGTGTCC 4089
Db 4193 GTTATATTTCTGCTGTGAGAACTTTAAAGTTTCAATACCAAGATGCTTCAGTGTCC 4252
Qy 4090 ATAGTGAATTTCTATGAGCCCAAGACAGCGCGGTGAGAGTTTCAACTCTGAAGTGAAG 4149
Db 4253 ATAGTGAATTTCTATGAGCCCAAGACAGCGCGGTGAGAGTTTCAACTCTGAAGTGAAG 4312
Qy 4150 CTGTCTCTCTGAGACCTTTGAGATGATGTCAGAGGCTGCGCTCTGTGAGAGTGAAGT 4209
Db 4313 CTGTCTCTCTGAGACCTTTGAGATGATGTCAGAGGCTGCGCTCTGTGAGAGTGAAGT 4372
Qy 4210 TCAGAGCTCCCATCATCACTCTTCAGTCAATTTTATTTTCTGTTCAAGCTTCTGACTTT 4269
Db 4373 TCAGAGCTCCCATCATCACTCTTCAGTCAATTTTATTTTCTGTTCAAGCTTCTGACTTT 4432
Qy 4270 ATGGAATCTTGGCTGTGA 4287
Db 4433 ATGGAATCTTGGCTGTGA 4450

RESULT 2
AF10459
LOCUS AF10459 5883 bp mRNA linear PRI 02-MAR-2002
DEFINITION Homo sapiens CD109 (CD109) mRNA, complete cds.
ACCESSION AF10459
VERSION AF10459.1 GI:19071208
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 5883)
Lin.M., Sutherland,D.R., Horsfall,W., Totty,N., Yeo,E., Nayar,R.,
Wu,X.F. and Schuh,A.C.
Cell surface antigen CD109 is a novel member of the alpha(2)
macroglobulin/C3, C4, C5 family of thioester-containing proteins
Blood 99 (5), 1693-1691 (2002)
JOURNAL MEDLINE
PUBMED 11861284
REFERENCE 2 (bases 1 to 5883)
AUTHORS Lin.M., Sutherland,D.R., Horsfall,W., Totty,N., Yeo,E., Nayar,R.,
Wu,X.F. and Schuh,A.C.
TITLE Direct Submission
JOURNAL Submitted (14-AUG-2001) Medicine, University of Toronto, 1 King's
College Circle, Room 7366, Toronto, Ontario M5S 1A8, Canada
FEATURES
source 1..5883
/organism="Homo sapiens"

```

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="K1"
1..5883
/gene="CD109"
1..112
/gene="CD109"
113..4450
/gene="CD109"
/note="associated with the Gov alloantigen system"
/codon_start=1
/product="CD109"
/protein_id="AAL84159.1"
/db_xref="GI:19071209"
/translation="MOGAPLTPAAHLICVCTALAAVAPGPRLPTAGLIRPGANVTI
GVELLHCPQVTVKAELEKTAENLTVSVLEAGVPEKSGFKTLTLPSPLEADREY
ELAVTERTODELIFNSSTRLEPTKRLISVLEQDPKALYKPKQVKEKRIYTLPEADREY
KISLNTLIDPKSNLIDQMSQSDLDGVLSTKIQLSHPILGWSIOVONDOTYOS
FOVSEYVLPEFVTLQTPLYCSNNSRLNGTITAKTYGKPVAGVTLFTPLPSFWGK
KNITITFKINGSNFSEFDEEMKNVMSDNGSEYLDLSSPPELITVTSVGI
SRVSTNVPEPKODYILIEPDYTVLKPSLNPATVYKTRADNOLTEERRNVVT
VPOBNTYVMSGNSGNOMKAEVOKINVTYPOSTGTEKIEPILDESSLELOKAYFIS
KSMAYHSLFKSPSTYITQKTRDENIKOSPELVVSGKRLKELSYRVSGQLVA
VGKQNSTMESELPENSWTPKACVIVYIEDEGELISDVLKI PVQLVFNKIKLYMSKY
KAPSEKSVSLRLISVTQPDSEIVGIVAVDKSVNLMANENDIMENVHLELNYGYLG
MFPNPSFAVPECEGLAVLTDLNLTQDYIDGIVDAEYDAERMEENEGHIVIDFSLGS
SPHVRKHPPEPTWTWIDNMGYRIYOBPEVTPPSTISWATGVISDGLGLTTPV
ELCARPPRTFLNLPYSVIRGEFALEITFNLYKQATVYKVIIEKSDKEDILMSE
INATGHOQTLVPSSEDAVTLFPIRTHLEIITITTTAALSPASDVAOTQMLKABEI
EKYSOSILDLIDNRLQSTLTKLTSFPENVTGSEVQVITAGVGLSGINGLASL
IRNPYCGDEGMNFPANFIYLDYLRKQKLTDLNLEKALSFRQGOBRLIYQREBG
SFSAFNGYDPSGSTMWLAFLVRCFLADPVIDIDQWLNHTYMLKQHSNCEPMDP
GRIYHSELQGNKSPVTLTAIVTSLGKRYKOPNDVOBSHFLSEFSGISDNYT
LALITVALSVSGSPKAEALNMLTWRAEDGCGMPVSSKXLSDSMOPSLIDVVA
YALSHFLOQTESEGPINRMLSRKNSLGRPSDQDTTALALSEFALNMTERN
IQVTVGSPSSPVKFLIDHNRLLIQTMELAVQPAVVISANGFEALCOLAVNL
VKAGSSRRRSRIONQEPDLDAVENKEDLHNVLDVSTSGRPSMALMEVNL
LSGFVPESEALISLSEYKVEYDGHKNTLIDVSENOFQVNI PAVANFVSTODAS
VSLVDYVEPRQAVRSYNSVVKLSODLGSQVDCRPECBDGASGSHHSVIFCFK
LTFPMLMT."
4451..5883
/gene="CD109"
4748
/misc_feature
/gene="CD109"
/note="alternative polyA site found on clone K1"
BASE COUNT 1697 a 1173 c 1241 g 1772 t
ORIGIN
Query Match 98.5%; Score 4224.4; DB 9; Length 5883;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 4286; Conservative 0; Mismatches 1; Indels 51; Gaps 1;
Qy 1 ATGAGGCGCCACCGCTCTGACCGCGCCCACTCTCTGCGTGTGACCGCGCGCTG 60
Db 113 ATGAGGCGCCACCGCTCTGACCGCGCCCACTCTCTGCGTGTGACCGCGCGCTG 172
Qy 61 GCGTGGCTCCCGGCGCTGCTTGTGAGACGCGCCAGAGATCATCAGCGCGAGGA 120
Db 173 GCGTGGCTCCCGGCGCTGCTTGTGAGACGCGCCAGAGATCATCAGCGCGAGGA 232
Qy 121 AATGTGACTATTTGGGAGTGTGAGACCTTGTGAAACAGTGCCTTCAGAGTGTGAAAGCG 180
Db 233 AATGTGACTATTTGGGAGTGTGAGACCTTGTGAAACAGTGCCTTCAGAGTGTGAAAGCG 292
Qy 181 GAGTGTCAAGACAGATCAAACTCTGCTGTCTGTCTGTGAGAGAGAGAGTCTTT 240
Db 293 GAGTGTCAAGACAGATCAAACTCTGCTGTCTGTCTGTGAGAGAGAGAGTCTTT 352
Qy 241 GAAAAAGCTCTTTTAAGACACTTCTTCATCACTACTCTGAGACAGTGCAGATGAG 300
Db 353 GAAAAAGCTCTTTTAAGACACTTCTTCATCACTACTCTGAGACAGTGCAGATGAG 412
Qy 301 ATTATAGCTACGTGTAACCGGAGTAACAGAGATGAGATTTATTTCTTAATAGTACC 360

```

Db 413 ATTATGACCTACGTGTAACCGGACGTACCAGAGATGAGATTTTATCTTAATATATACC 472  
Oy 361 CGCTTATCTTTGAGACCAAGAAATATCTGCTTCAATCAAAACAGACAGAGGCTTAATAC 420  
Db 473 CGCTTATCTTTGAGACCAAGAAATATCTGCTTCAATCAAAACAGACAGAGGCTTAATAC 532  
Oy 421 AAGCCAAAGCAAGAGTGAAGTTTGGCATTTGTACACTCTTCTCAGATTTTAAAGCTTAC 480  
Db 533 AAGCCAAAGCAAGAGTGAAGTTTGGCATTTGTACACTCTTCTCAGATTTTAAAGCTTAC 592  
Oy 481 AAAACCTCTTAAACATTTCTCAATTAAGAACCCCAATCAATTTGATCCAAGTGTGG 540  
Db 593 AAAACCTCTTAAACATTTCTCAATTAAGAACCCCAATCAATTTGATCCAAGTGTGG 652  
Oy 541 TCACAAACAAGTATCTTGAAGTCAATTTCCAAAATTTTCAGTATCTTCCATCCAAAT 600  
Db 653 TCACAAACAAGTATCTTGAAGTCAATTTCCAAAATTTTCAGTATCTTCCATCCAAAT 712  
Oy 601 GTTGTGACTGCTATTTCAAGTTCAGTGAATGACACAGACATATATCAATCATTTTCAG 660  
Db 713 GTTGTGACTGCTATTTCAAGTTCAGTGAATGACACAGACATATATCAATCATTTTCAG 772  
Oy 661 GTTTCAGATATGTATTTACCAAAATTTGAAGTCACTTTCAGACACCATTAATATGTTCT 720  
Db 773 GTTTCAGATATGTATTTACCAAAATTTGAAGTCACTTTCAGACACCATTAATATGTTCT 832  
Oy 721 ATGAATTCAGATTTAAATGATGACCATACGGGCAAGATATACATATGAGGAAGCCAGTG 780  
Db 833 ATGAATTCAGATTTAAATGATGACCATACGGGCAAGATATACATATGAGGAAGCCAGTG 892  
Oy 781 AAAGAGACGTAAACGCTTACATTTTACCTTTATCCTTTTGGGGAAGAAATATAT 840  
Db 893 AAAGAGACGTAAACGCTTACATTTTACCTTTATCCTTTTGGGGAAGAAATATAT 952  
Oy 841 ACAAACCAATTTAAGATAATGATCTGCAAACTTCTCTTTTAAATGATGAAGATGAA 900  
Db 953 ACAAACCAATTTAAGATAATGATCTGCAAACTTCTCTTTTAAATGATGAAGATGAA 1012  
Oy 901 AATGTATGATGATCTTCAATGAGACTTTCGAAATACCTGATCTATCTTCCCTGAGACA 960  
Db 1013 AATGTATGATGATCTTCAATGAGACTTTCGAAATACCTGATCTATCTTCCCTGAGACA 1072  
Oy 961 GTAGAATTTTAAACCAAGTGAACAGATCAGTTACAGTATTTCAAGAATGTAAAGCAT 1020  
Db 1073 GTAGAATTTTAAACCAAGTGAACAGATCAGTTACAGTATTTCAAGAATGTAAAGCAT 1132  
Oy 1021 AATGTGTTCTTCAAGACATGATTAATCATATGAGTTTGTATTATACTACTGTCTTG 1080  
Db 1133 AATGTGTTCTTCAAGACATGATTAATCATATGAGTTTGTATTATACTACTGTCTTG 1192  
Oy 1081 AAGCCATCTCTCAACTTCACAGCACTGAGAAAGTAACTGCTGATGAGCAACCACTG 1140  
Db 1193 AAGCCATCTCTCAACTTCACAGCACTGAGAAAGTAACTGCTGATGAGCAACCACTG 1252  
Oy 1141 ACTCTTGAAGAAAGAAATATATGATGATGATGATGATGATGATGATGATGATGATG 1200  
Db 1253 ACTCTTGAAGAAAGAAATATATGATGATGATGATGATGATGATGATGATGATGATG 1312  
Oy 1201 TACTGAGAGGATCTTAACAGTGAATTCAGAAAATGGAAGCTGTTCAAGAAAATTAAT 1260  
Db 1313 TACTGAGAGGATCTTAACAGTGAATTCAGAAAATGGAAGCTGTTCAAGAAAATTAAT 1372  
Oy 1261 ACTGTCCCCCAAGTGAAGCTTTTAAGATTTGAATCTCCCAATCTGAGAGATTCAGTGG 1320  
Db 1373 ACTGTCCCCCAAGTGAAGCTTTTAAGATTTGAATCTCCCAATCTGAGAGATTCAGTGG 1432  
Oy 1321 CTACAGTTGAAGGCTATTTCTTGTAGTAAAGATGAGTGAAGTGAAGTGAAGTGAAG 1380  
Db 1433 CTACAGTTGAAGGCTATTTCTTGTAGTAAAGATGAGTGAAGTGAAGTGAAGTGAAG 1492  
Oy 1381 AAGTCTCTAGTAAAGCATATCATCAATTAAGAAACAGAGATGAATTAAGTGAAG 1440

Db 1493 AAGTCTCTAGTAAAGCATATCATCAACTTAAGAAACAGAGATGAATTAAGTGAAG 1552  
Oy 1441 TCGCCTTTTGAAGTGTGTAGTGGCAACAAACATTAAGAGATTAAGTGAATTAAGTGA 1500  
Db 1553 TCGCCTTTTGAAGTGTGTAGTGGCAACAAACATTAAGAGATTAAGTGAATTAAGTGA 1612  
Oy 1501 GTATCAGAGGACAGTGTGTGTAGTGGCAACAAACATTAAGAGATTAAGTGAATTAAG 1560  
Db 1613 GTATCAGAGGACAGTGTGTGTAGTGGCAACAAACATTAAGAGATTAAGTGAATTAAG 1672  
Oy 1561 CCAGAAATTTCTTGAAGCTTCCAAAACCTGTGTAAATTTGTATTAATTAAGAGATGA 1620  
Db 1673 CCAGAAATTTCTTGAAGCTTCCAAAACCTGTGTAAATTTGTATTAATTAAGAGATGA 1732  
Oy 1621 GAATTTAATGAAGTGTCTTAATAAATTTCTGTTCAGTGTGTTTAAATTAAGTGAAG 1680  
Db 1733 GAATTTAATGAAGTGTCTTAATAAATTTCTGTTCAGTGTGTTTAAATTAAGTGAAG 1792  
Oy 1681 CTATATTTGAAGTAAAGTGAAGCTGAAACCATCTGAGAAAGTCTCTTGAAGATCTGTG 1740  
Db 1793 CTATATTTGAAGTAAAGTGAAGCTGAAACCATCTGAGAAAGTCTCTTGAAGATCTGTG 1852  
Oy 1741 ACACAGCTGACTCATATGTTGGATTTGATCTGTGCAAAAGTGTGAATCTGATGAAT 1800  
Db 1853 ACACAGCTGACTCATATGTTGGATTTGATCTGTGCAAAAGTGTGAATCTGATGAAT 1912  
Oy 1801 GCCTTAATGATATTAAGTGAAGAAATGAGTGCATGATGTTGAAGCTTTATPAACAGAG 1860  
Db 1913 GCCTTAATGATATTAAGTGAAGAAATGAGTGCATGATGTTGAAGCTTTATPAACAGAG 1972  
Oy 1861 TATTAATTTAGGATGTCATGAATTTCTTTTCAGTCTTTTCAGAAATGTGACTCGGGTA 1920  
Db 1973 TATTAATTTAGGATGTCATGAATTTCTTTTCAGTCTTTTCAGAAATGTGACTCGGGTA 2032  
Oy 1921 TTGACAGATGCAAACTCTCAAGAGATTAATGATGTTTATGACATGACAGATAT 1980  
Db 2033 TTGACAGATGCAAACTCTCAAGAGATTAATGATGTTTATGACATGACAGATAT 2092  
Oy 1981 GCTGAGAGTTTATGAGAGAAATGAGACATTTGATGATTAATCATGACTTTCTTGG 2040  
Db 2093 GCTGAGAGTTTATGAGAGAAATGAGACATTTGATGATTAATCATGACTTTCTTGG 2152  
Oy 2041 GGTACAGTTCACATGTCGAAAGATTTTCAGAGACTTTCAGAGATTTGCTAGACACCAAC 2100  
Db 2153 GGTACAGTTCACATGTCGAAAGATTTTCAGAGACTTTCAGAGATTTGCTAGACACCAAC 2212  
Oy 2101 ATGGGTTACAGGATTTACCAAGATTTGAAGTAACTGATCTGATTTATCACTTCTGG 2160  
Db 2213 ATGGGTTACAGGATTTACCAAGATTTGAAGTAACTGATCTGATTTATCACTTCTGG 2272  
Oy 2161 GTGGTACTGTGTTTGTGATCTCTGAGAGACTTGGGCTTGGACATPAACATTAATCTCAAG 2220  
Db 2273 GTGGTACTGTGTTTGTGATCTCTGAGAGACTTGGGCTTGGACATPAACATTAATCTCAAG 2332  
Oy 2221 GAGCTCCAAAGCTTCCAAACATTTTTCATTTTGTGAATCTTCCCTCATCTGTTATCGA 2280  
Db 2333 GAGCTCCAAAGCTTCCAAACATTTTTCATTTTGTGAATCTTCCCTCATCTGTTATCGA 2392  
Oy 2281 GGTGAAGATTTGCTTGTGAATTAATCTATATTAATTAATTAAGATGCACTGAGGTT 2340  
Db 2393 GGTGAAGATTTGCTTGTGAATTAATCTATATTAATTAATTAAGATGCACTGAGGTT 2452  
Oy 2341 AAGGTATCATTTGAGAAAGTGAACAAATTTGATATCTTAAGATCTTCAAGTGAATTAAT 2400  
Db 2453 AAGGTATCATTTGAGAAAGTGAACAAATTTGATATCTTAAGATCTTCAAGTGAATTAAT 2512  
Oy 2401 GCCACAGGCAACAGCAGACACCTTCTGTGTTCCAGTGAAGATGAGGCAAGTCTTCTTTT 2460  
Db 2513 GCCACAGGCAACAGCAGACACCTTCTGTGTTCCAGTGAAGATGAGGCAAGTCTTCTTTT 2572  
Oy 2461 CCCATCAGGCCAACAATCTGTGGAGAAATTTCTATCAAGTCAAGCTCTTTCACCACT 2520  
Db 2573 CCCATCAGGCCAACAATCTGTGGAGAAATTTCTATCAAGTCAAGCTCTTTCACCACT 2632



QY 2521 GCTTCTGATGCTGTCACCCAGATGATTTTAAAGGCTGAAGAAATAGAAAAATCATAT 2580  
 DB 2633 GCTTCTGATGCTGTCACCCAGATGATTTTAAAGGCTGAAGAAATAGAAAAATCATAT 2692  
 QY 2581 TCACATCATCTTATTTAGACTGACGACATAGGCTGACAGAGTACCTGAAAACTTTG 2640  
 DB 2693 TCACATCATCTTATTTAGACTGACGACATAGGCTGACAGAGTACCTGAAAACTTTG 2752  
 QY 2641 AGTTCTCATCTTCCCTTAATAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2700  
 DB 2753 AGTTCTCATCTTCCCTTAATAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2812  
 QY 2701 GGAAGTGTCTTGTGCTTCCATCAATGCTTAAGCTTCAATGATTCGATCCCTTAATGCG 2760  
 DB 2813 GGAAGTGTCTTGTGCTTCCATCAATGCTTAAGCTTCAATGATTCGATCCCTTAATGCG 2872  
 QY 2761 TGTGTGTAACAGAACATGATTAATTTGCTCCAAATATTTACATTTGGATTTATCTGACT 2820  
 DB 2873 TGTGTGTAACAGAACATGATTAATTTGCTCCAAATATTTACATTTGGATTTATCTGACT 2932  
 QY 2821 AAAAAAGAACATGACGATTAATTTGAAGAAAGCTTTCAATTAATAGGCAAGCT 2880  
 DB 2933 AAAAAAGAACATGACGATTAATTTGAAGAAAGCTTTCAATTAATAGGCAAGCT 2992  
 QY 2881 TACAGAGAGAACTTCTCTATCAGAGGAGATGCTCTTCAATGCTTTGGAAATTTAT 2940  
 DB 2993 TACAGAGAGAACTTCTCTATCAGAGGAGATGCTCTTCAATGCTTTGGAAATTTAT 3052  
 QY 2941 GACCTTCTGAGAGACTTGTGCTGCTGCTTTGTTTAAATGATTTCTTTGAAGCCGAT 3000  
 DB 3053 GACCTTCTGAGAGACTTGTGCTGCTGCTTTGTTTAAATGATTTCTTTGAAGCCGAT 3112  
 QY 3001 CCTTACATAGATTTGATCAGAAATGCTTTACAGAAATACATCTTGGCTTTAAAGCAAT 3060  
 DB 3113 CCTTACATAGATTTGATCAGAAATGCTTTACAGAAATACATCTTGGCTTTAAAGCAAT 3172  
 QY 3061 CAGAAATCAACGCTGATTTTGGATCCAGAAAGATGATTTCAATGAGACTTCAAGGT 3120  
 DB 3173 CAGAAATCAACGCTGATTTTGGATCCAGAAAGATGATTTCAATGAGACTTCAAGGT 3232  
 QY 3121 GGCATTAATAAGTCCAGTAACCTTACAGCTTAATTTGTAATTTCTCTCGGAAATTAAGA 3180  
 DB 3233 GGCATTAATAAGTCCAGTAACCTTACAGCTTAATTTGTAATTTCTCTCGGAAATTAAGA 3292  
 QY 3181 AAGATACGCTTACATTTGATGTCAGAGATGCTTCAATTTTGGAGTCCGATTTCACT 3240  
 DB 3293 AAGATACGCTTACATTTGATGTCAGAGATGCTTCAATTTTGGAGTCCGATTTCACT 3352  
 QY 3241 AGAGAAATTTGAGCAATTTACTCTAGCCCTTAATTTGATTTGATTTGATTTGATTTGAT 3300  
 DB 3353 AGAGAAATTTGAGCAATTTACTCTAGCCCTTAATTTGATTTGATTTGATTTGATTTGAT 3412  
 QY 3301 AGTCTTAAGGGAAGAAAGCTTTGATTTGCTGACTTGAAGAGAGAAAGAAAGTGGC 3360  
 DB 3413 AGTCTTAAGGGAAGAAAGCTTTGATTTGCTGACTTGAAGAGAGAAAGAAAGTGGC 3472  
 QY 3361 ATGCAATTTCTGGGTGCTATCAGAGTCCAAATTTTCTGACTCTTGGAGCCAGCCCTTCT 3420  
 DB 3473 ATGCAATTTCTGGGTGCTATCAGAGTCCAAATTTTCTGACTCTTGGAGCCAGCCCTTCT 3532  
 QY 3421 GATATTTGAAGTTGAGCTTATGACCTGCTTCAACCTTTTACAAATTTTCAAGCTTTGAG 3480  
 DB 3533 GATATTTGAAGTTGAGCTTATGACCTGCTTCAACCTTTTACAAATTTTCAAGCTTTGAG 3592  
 QY 3481 GGAATCCCAATTTATAGGTGCTTAAGAGGCAAGAAATAGCTTGGTGTGTTTGCATCT 3540  
 DB 3593 GGAATCCCAATTTATAGGTGCTTAAGAGGCAAGAAATAGCTTGGTGTGTTTGCATCT 3652  
 QY 3541 ACTCAGGATACCACTGTGGCTTTAAAGGCTTGTCTGATTTTGAAGCCCTTAATTAAGA 3600  
 DB 3653 ACTCAGGATACCACTGTGGCTTTAAAGGCTTGTCTGATTTTGAAGCCCTTAATTAAGA 3712

QY 3601 GAAAGACAAATATTCAGATGACCGGTGACCGGCTTACCTCACAACTCTT----- 3651  
 DB 3713 GAAAGACAAATATTCAGATGACCGGTGACCGGCTTACCTCACAACTCTTAAAGTTT 3772  
 QY 3652 -----CTTGCTGTGTAAGCCCA 3669  
 DB 3773 CTGATTTGACACACACACCGCTTATCTCTTCAAGACAGACAGAGCTGTGTGTAAGCCCA 3832  
 QY 3670 ATGGCAGTTAATATTTCCGCAATAGTTTGTGATTTGCTATTTGTCAGCTCAATGTTGTA 3729  
 DB 3833 ATGGCAGTTAATATTTCCGCAATAGTTTGTGATTTGCTATTTGTCAGCTCAATGTTGTA 3892  
 QY 3730 TATATGTAAGGCTTCTGGCTTCTTGAAGACGAAGATCTATCCAAATCAAGAAAGCC 3789  
 DB 3893 TATATGTAAGGCTTCTGGCTTCTTGAAGACGAAGATCTATCCAAATCAAGAAAGCC 3952  
 QY 3790 TTTGATTTAATGTTGCTGTAAAGAAATTAAGATGATCTCAATGATGTAATTTGAAT 3849  
 DB 3953 TTTGATTTAATGTTGCTGTAAAGAAATTAAGATGATCTCAATGATGTAATTTGAAT 4012  
 QY 3850 GTGTGTCAAGCTTTTCCGCGCCGCGGTGAGAGTGGCCTTATGGAAGTTAACTTA 3909  
 DB 4013 GTGTGTCAAGCTTTTCCGCGCCGCGGTGAGAGTGGCCTTATGGAAGTTAACTTA 4072  
 QY 3910 TTAAGTGGCTTTATGCTGCTTCAAGAAATTTCTTGAAGACGACATGAAAGAAAGTG 3969  
 DB 4073 TTAAGTGGCTTTATGCTGCTTCAAGAAATTTCTTGAAGACGACATGAAAGAAAGTG 4132  
 QY 3970 GAATATGATCATGGAAGAACTCAACCTTATTTGATTTCTGTAATGAAACCCAGTTTGT 4029  
 DB 4133 GAATATGATCATGGAAGAACTCAACCTTATTTGATTTCTGTAATGAAACCCAGTTTGT 4192  
 QY 4030 GTTAATTTCTGCTGTGAGAACTTTAAAGTTTCAATCCCAAGATGCTTCAAGTCTC 4089  
 DB 4193 GTTAATTTCTGCTGTGAGAACTTTAAAGTTTCAATCCCAAGATGCTTCAAGTCTC 4252  
 QY 4090 ATAGTGAATTTACTATGAGCCCAAGAGACAGCGGTGAGAAATTTCAACTTGAAGTGAAG 4149  
 DB 4253 ATAGTGAATTTACTATGAGCCCAAGAGACAGCGGTGAGAAATTTCAACTTGAAGTGAAG 4312  
 QY 4150 CTGTCTCTCTGTGACCTTTGCAATGATGTCAGAGGCTGCGCTTGTGAGATGAGACT 4209  
 DB 4313 CTGTCTCTCTGTGACCTTTGCAATGATGTCAGAGGCTGCGCTTGTGAGATGAGACT 4372  
 QY 4210 TCAAGCTCCCATGATCACTTTCAGTCAATTTTATTTTGTTCAGAGCTTCTGTAATTT 4269  
 DB 4373 TCAAGCTCCCATGATCACTTTCAGTCAATTTTATTTTGTTCAGAGCTTCTGTAATTT 4432  
 QY 4270 ATGAACTTTGGCTGTGA 4287  
 DB 4433 ATGAACTTTGGCTGTGA 4450

RESULT 3  
 AX534938  
 LOCUS 5895 bp DNA linear PAT 22-NOV-2002  
 DEFINITION Sequence 5 from Patent WO02070696.  
 ACCESSION AX534938  
 VERSION AX534938.1 GI:25261482  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE  
 1 Schuh, A. and Sutherland, R.D.  
 CD109 nucleic acid molecules polypeptides and methods of use  
 Patent: WO 02070696-A 5 12-SEP-2002;  
 Schuh, Andrew Sutherland, Robert D. (CA)  
 Location/Qualifiers  
 1..5895  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"

same as 1



CDS

/db\_xref="taxon:9606"  
113..4450  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD57233.1"  
/db\_xref="GI:25261483"  
/translation="MQGPPPLTRAHLLCVCTAALAVAPGPRFLVAPGIIIRGNVTT  
GELLEHCPSQVFAELKLTASNLTVSLAEAGVEKSFETLLPSLPLMSADEIY  
ELRVGTODEILFNSRSLSEFKRIYFIDTKALYPRKQEVRIYLFSDRPY  
KSLNLIIDPKSNLIQWMLSOOSDLGSKTFOLSHPIIDGWSLOYVNOQTYOS  
FOVSEYLPKPEVTLQTLPLCYGMSKHNGLITAKTYKPKVGVTLTLESTMG  
KNITKPEKINGSANPSFDEMKMNSNGLSEYDSSGPRPEILITTESYTG  
SRVSTNVFKQHDIIIEFFDYTVLKLPSLNTFAIVKTRAGNOQLTEERNNAVIT  
VTQRYNTEYWSGNSGNQMEAVQKINYPQSGTFKLEPILLESSELOKSGEYF  
KSMNAVHSLFKSPSTYIOLKTRDENIKYSPFELVANSKELKEISYLVVARGDLA  
VGKNSMTSLPENSMTKACIVYIYIDDEGIIEITVIALSPASADVOMILVKAEG  
KAPSEKSLRI SVTOPDSIVGIIVADVKNIMASNDITMENVAHELEMYGYG  
MEMNSFAYROEGWVLTPLNLTKEYIIDVYDINAEAEPRMEENSHYDIDFSLGS  
SFVARKHPEPFIWLDITNGYRIYQEFVTVDSITSWATFVISEDGLDITTPV  
ELQARQPFPIFLNLPYSVIRGEFALEITIENYLDATVVKYIIEKSDKFIIMTNS  
INATGHOQLLVSEDEGATVLEPIRPHLGEIPIVTVALSPPASADVOMILVKAEG  
EKSYOSIILDLTNRLOSTLKTLSFPSPNTVTSERVOITAGIDVILGSLASL  
IMPYGCGEOMINIPAPNTYIIDYLTKKQOLDNKKERLSEMGQYQRELLYQEDG  
SFSAQNTYDPSGSTWLSAFVLCFLEADYIDIDQVYLRHTTWLKGKSGEYFDP  
GVVHSELQGNKSPVTLTAIVLISLGRKYQPNIDVQESIHFLSESESRISDNTY  
IALITYALSVSQPRKALNMLTWRAEGEGQFVWSSESLSDSMOPRSIDIEVA  
VALSHLOFQTSSEGIPIRMWLSRONSLEGFASQDITVALKALSEFALNTRTN  
IOVTVGSPSPKVEKIDITHRLLQTELAIVOPMAVNIANGGFAICOLNVVN  
VAKSSRRRSRISIOAEAPDLDAVAKENDDLNHDINDCTSFSPGRSGMALMEVNL  
LSGFWYPSBAISLSETVKVEYDHEKLNLYDSVNETQCVNIPAVRNKVSNTDAS  
VSIIVDYERQRAVRSYNSVRLSSCDLSDVQGRPCBDASGSHHSVIFICFK  
LYFMELMWL"

BASE COUNT 1709 a 1173 c 1241 g 1772 c

ORIGIN

Query Match 98.5%; Score 4224.4; DB 6; Length 5895;  
Best Local Similarity 98.8%; Pred. No. 0;  
Matches 4286; Conservative 0; Mismatches 1; Indels 51; Gaps 1;

1 ATGCAAGGCGCCACCGCTCCGACCGCGCGCCACCTCTCGCGTGGACCGCGCGCGC 60  
113 ATGCAAGGCGCCACCGCTCCGACCGCGCGCCACCTCTCTGCGTGGACCGCGCGCGC 172  
61 GCCGTGGCTCCGCGCGCTCGTTCGTGACAGCCCGCAGGATCATCAGCCCGGAGA 120  
173 GCGTGGCTCCGCGCGCTCGTTCGTGACAGCCCGCAGGATCATCAGCCCGGAGA 232  
121 AATGTACTATTGGGGTGGAGCTTCTGAAACCTGCGCTTCAAGGTGACTGTGAAGGG 180  
233 AATGTACTATTGGGGTGGAGCTTCTGAAACCTGCGCTTCAAGGTGACTGTGAAGGG 292  
181 GAGCTGCTAAGACAGCATCAAACTCATCTGTCTGTCTGGAAGCAGAAGAGCTTT 240  
293 GAGCTGCTAAGACAGCATCAAACTCATCTGTCTGTCTGGAAGCAGAAGAGCTTT 352  
241 GAAAAAGCTCTTTAAGACACTTACTCTTCCATCACTACTCTGAAAGCTGGAAGTAG 300  
353 GAAAAAGCTCTTTAAGACACTTACTCTTCCATCACTACTCTGAAAGCTGGAAGTAG 412  
301 ATTATGACCTACGTGTAAACCGACGTACCCAGAGATGAGATTTATTTCTAATAGTACC 360  
413 ATTATGACCTACGTGTAAACCGACGTACCCAGAGATGAGATTTATTTCTAATAGTACC 472  
361 CGCTATCATTTGAGACCAAGAGAAATATGTCTTCAATCAACAGACAGAGCTTATAC 420  
473 CGCTATCATTTGAGACCAAGAGAAATATGTCTTCAATCAACAGACAGAGCTTATAC 532  
421 AAGCCAAAGCAAGAGAGTTCGATTTGATCACTCTCTCAGATTTTAAGCTTAC 480  
533 AAGCCAAAGCAAGAGAGTTCGATTTGATCACTCTCTCAGATTTTAAGCTTAC 592  
481 AAAACCTCTTTAAACATTTCTATTAAAGACCCCAAAATCAATTTGATTCACAGTGTG 540

593 AAAACCTCTTTAAACATTTCTATTAAAGACCCCAAAATCAAAATTTGATTCACAGTGTG 652  
541 TCACAAAGCAAGAGATCTTGAAGTCAATTTCAAAACCTTTACAGTATCTTCCATCAATA 600  
653 TCACAAAGCAAGAGATCTTGAAGTCAATTTCAAAACCTTTACAGTATCTTCCATCAATA 712  
601 CTTGCTGACGTGCTATTCAAGTTCAAGTGAATGACAGACATATTTATCAATCATTTGAG 660  
713 CTTGCTGACGTGCTATTCAAGTTCAAGTGAATGACAGACATATTTATCAATCATTTGAG 772  
661 GTTTAGAAATATATGTTTACCAAAATTTGAAGTCAATTTGACAGACATATTTATTTCT 720  
773 GTTTAGAAATATATGTTTACCAAAATTTGAAGTCAATTTGACAGACATATTTATTTCT 832  
721 ATGAATTTCAAGCATTTTAAATGTTGTCATCAAGCAAGTATATCATATGGAAGCAGTG 780  
833 ATGAATTTCAAGCATTTTAAATGTTGTCATCAAGCAAGTATATCATATGGAAGCAGTG 892  
781 AAAGAGACGTAAACGCTTACATTTTACCTTTATCTTTTGGGGAAGAGAAAATATTT 840  
893 AAAGAGACGTAAACGCTTACATTTTACCTTTATCTTTTGGGGAAGAGAAAATATTT 952  
841 AAAAAAATTTAAGATTAATGATTAATGATCTGCAAACTCTCTTTAATGATGAAGATGAA 900  
953 AAAAAAATTTAAGATTAATGATTAATGATCTGCAAACTCTCTTTAATGATGAAGATGAA 1012  
901 AATGTAATGATATTTCTCAATGAGACTTCTGTAATCTGTAATCTATCTTCCCTGAGACA 960  
1013 AATGTAATGATATTTCTCAATGAGACTTCTGTAATCTGTAATCTATCTTCCCTGAGACA 1072  
961 GTGAATATTTTAAACCAAGTGAACAGAAATCAGTTACAGGATTTTCAAGAAATGAAGACT 1020  
1073 GTGAATATTTTAAACCAAGTGAACAGAAATCAGTTACAGGATTTTCAAGAAATGAAGACT 1132  
1021 AATGTTCTTCAAGCAATGATTAATCATGATGTTTGTGATTAATCATCTGCTTG 1080  
1133 AATGTTCTTCAAGCAATGATTAATCATGATGTTTGTGATTAATCATCTGCTTG 1192  
1081 AAGCATCTCTCACTTCAAGCAATGATTAATCATGATGTTTGTGATTAATCATCTGCTTG 1140  
1193 AAGCATCTCTCACTTCAAGCAATGATTAATCATGATGTTTGTGATTAATCATCTGCTTG 1252  
1141 ACTCTTGAAGAAAGAAATTAATGATTAATGATTAATGATTAATGATTAATGATTAATG 1200  
1253 ACTCTTGAAGAAAGAAATTAATGATTAATGATTAATGATTAATGATTAATGATTAATG 1312  
1201 TACTGAGCGGATCTTAACAGTGAAGAAATGAGAAATGAGAAATGAGAAATGAGAAATG 1260  
1313 TACTGAGCGGATCTTAACAGTGAAGAAATGAGAAATGAGAAATGAGAAATGAGAAATG 1372  
1261 ACTGCCCCCAAGTGAAGAAATTTAAGTTGAATTCCTCAATCTGAGAGATTTCCAGTAG 1320  
1373 ACTGCCCCCAAGTGAAGAAATTTAAGTTGAATTCCTCAATCTGAGAGATTTCCAGTAG 1432  
1321 CTACAGTTGAAGGCTTATTTCTTGTGTAAGTAAAGTACATGAGCTCATAGTCTGTT 1380  
1433 CTACAGTTGAAGGCTTATTTCTTGTGTAAGTAAAGTACATGAGCTCATAGTCTGTT 1492  
1381 AAGTCTCTTAAGACATCAATCACTAAACCAAGAGTGAATATTAAGTGAAGTGAAGTGA 1440  
1493 AAGTCTCTTAAGACATCAATCACTAAACCAAGAGTGAATATTAAGTGAAGTGAAGTGA 1552  
1441 TCGCTTTTGAAGTGTGTGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1500  
1553 TCGCTTTTGAAGTGTGTGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1612  
1501 GTATCCAGGAGACAGTGTGTGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1560  
1613 GTATCCAGGAGACAGTGTGTGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1672  
1561 CCAGAAAATTTTGGACTCCAAAAGCTGTGTAAATGTGTAAATTTGATTAATGAAGATGAGG 1620  
1673 CCAGAAAATTTTGGACTCCAAAAGCTGTGTAAATGTGTAAATTTGATTAATGAAGATGAGG 1732

QY 1621 GAAATTATAGATGTTCTAAAAATTCCTGTCAGCTTGTTTAAAAATPAAGATAAG 1680  
Db 1723 GAAATTATAGATGTTCTAAAAATTCCTGTCAGCTTGTTTAAAAATPAAGATAAG 1792  
QY 1681 CTATATTTGGAGTAAAGTGAAGTGAACCATCTGAGAAAGTCTCTTTAGATCTCTGTG 1740  
Db 1793 CTATATTTGGAGTAAAGTGAAGTGAACCATCTGAGAAAGTCTCTTTAGATCTCTGTG 1852  
QY 1741 ACACAGCTGATCCATAGTTGGATTTGAGTGTGACAAAGTGTGAATCTGATGAT 1800  
Db 1853 ACACAGCTGATCCATAGTTGGATTTGAGTGTGACAAAGTGTGAATCTGATGAT 1912  
QY 1801 GCTCTTAATGATATTACAAATGAAAATGTGTCCATGAGTTGAACTTTATPACACAGA 1860  
Db 1913 GCTCTTAATGATATTACAAATGAAAATGTGTCCATGAGTTGAACTTTATPACACAGA 1972  
QY 1861 TATATTTAGGCAATGTTATGATATTTCTTGACAGTCTTCAGAAATGTGACTGGGTA 1920  
Db 1973 TATATTTAGGCAATGTTATGATATTTCTTGACAGTCTTCAGAAATGTGACTGGGTA 2032  
QY 1921 TTGACAGATGCAAACTCAGCAAGGATTAATATGATGTTTATGACAAATGCAAGATAT 1980  
Db 2033 TTGACAGATGCAAACTCAGCAAGGATTAATATGATGTTTATGACAAATGCAAGATAT 2092  
QY 1981 GCTGAGAGTTTATGAGAAAATGAAAGACATATTTGATATATTCATGACTTTTCTTTG 2040  
Db 2093 GCTGAGAGTTTATGAGAAAATGAAAGACATATTTGATATATTCATGACTTTTCTTTG 2152  
QY 2041 GGTAGCAATCCATGTCGAAAGATTTTCCAGAGACTTGGATTTGGCTGACACCAAC 2100  
Db 2153 GGTAGCAATCCATGTCGAAAGATTTTCCAGAGACTTGGATTTGGCTGACACCAAC 2212  
QY 2101 ATGGATTACAGATTTTACCAAGATTTGAGTAATCTGATCTGATCTGATCTTTG 2160  
Db 2213 ATGGATTACAGATTTTACCAAGATTTGAGTAATCTGATCTGATCTGATCTTTG 2272  
QY 2161 GTGGCTACTGTTTGTGATCTGAGAGACCTGGCTCTTGACATPACAACTACTCCAGTG 2220  
Db 2273 GTGGCTACTGTTTGTGATCTGAGAGACCTGGCTCTTGACATPACAACTACTCCAGTG 2332  
QY 2221 GAGCTCCAGCTTCCAAACATTTTCAATTTTGTGAATCTTCCCTACTCTGTTATCGA 2280  
Db 2333 GAGCTCCAGCTTCCAAACATTTTCAATTTTGTGAATCTTCCCTACTCTGTTATCGA 2392  
QY 2281 GGTGAAGATTTGCTTGGAAATACTATATTTCAATTTTGAAGATGCCAGAGGT 2340  
Db 2393 GGTGAAGATTTGCTTGGAAATACTATATTTCAATTTTGAAGATGCCAGAGGT 2452  
QY 2341 AAGGTAAATCATTTGAGAAAAGTGAACAATTTGATATTTCAATGACTTCAAGTAATAAT 2400  
Db 2453 AAGGTAAATCATTTGAGAAAAGTGAACAATTTGATATTTCAATGACTTCAAGTAATAAT 2512  
QY 2401 GCCACAGGCCACACAGAGACCTTCTGTGTTCCAGTGAAGATGGGGCAACTGTTCTTTT 2460  
Db 2513 GCCACAGGCCACACAGAGACCTTCTGTGTTCCAGTGAAGATGGGGCAACTGTTCTTTT 2572  
QY 2461 CCCATCAGGCCAACAACATCTGGAGAAATTCCTATCAAGTCAAGCTCTTCAACCACT 2520  
Db 2573 CCCATCAGGCCAACAACATCTGGAGAAATTCCTATCAAGTCAAGCTCTTCAACCACT 2632  
QY 2521 GCTTCTGATGCTGTCACCCAGATGATTTTAAAGGCTGAAGAAATGAAAATCTATAT 2580  
Db 2633 GCTTCTGATGCTGTCACCCAGATGATTTTAAAGGCTGAAGAAATGAAAATCTATAT 2692  
QY 2581 TCACAAATCCATTTATAGACTGACTGACAAATAGGCTACAGAGTACCTGAAAATCTTG 2640  
Db 2693 TCACAAATCCATTTATAGACTGACTGACAAATAGGCTACAGAGTACCTGAAAATCTTG 2752  
QY 2641 AGTTTCATTTCTCCCTAATATACAGTGACTGGCAGTGAAGAAAGTTGCAATCTGCAAT 2700  
Db 2753 AGTTTCATTTCTCCCTAATATACAGTGACTGGCAGTGAAGAAAGTTGCAATCTGCAAT 2812

QY 2701 GGAGATGTTCTTGCTCTTCATCATATGCTTACGCTTACGCTCATTTGATGCGATGCTTATGCG 2760  
Db 2813 GGAGATGTTCTTGCTCTTCATCATATGCTTACGCTTACGCTCATTTGATGCGATGCTTATGCG 2872  
QY 2761 TGTGCTAACAAGAACATGATTAATTTGCTCCAAATTTTATCAATTTGATGATCTGACT 2820  
Db 2873 TGTGCTAACAAGAACATGATTAATTTGCTCCAAATTTTATCAATTTGATGATCTGACT 2932  
QY 2821 AAAAAGAAACAATGACAGATTAATTTGAAAAGAAAAGCTCTTTCATTTATGAGCAAGT 2880  
Db 2933 AAAAAGAAACAATGACAGATTAATTTGAAAAGAAAAGCTCTTTCATTTATGAGCAAGT 2992  
QY 2881 TACCAAGAGAACTTCTCTATACAGAGGAAGATGCTCTTCACTGCTTTTGGAAATTA 2940  
Db 2993 TACCAAGAGAACTTCTCTATACAGAGGAAGATGCTCTTCACTGCTTTTGGAAATTA 3052  
QY 2941 GACCTTCTGGAGCACTTGCTGTCAGCTTTGTTTAAAGATGTTCTTGAAGCCGAT 3000  
Db 3053 GACCTTCTGGAGCACTTGCTGTCAGCTTTGTTTAAAGATGTTCTTGAAGCCGAT 3112  
QY 3001 CCTTACATAGATATTTGATCAGAAATGTTTACAGAAACATACACTTGGCTTAAAGACAT 3060  
Db 3113 CCTTACATAGATATTTGATCAGAAATGTTTACAGAAACATACACTTGGCTTAAAGACAT 3172  
QY 3061 CAGAAATCCAAAGCTGAATTTTGGATCCAGAAAGATGATTCATGATGACTTCAAGGT 3120  
Db 3173 CAGAAATCCAAAGCTGAATTTTGGATCCAGAAAGATGATTCATGATGACTTCAAGGT 3232  
QY 3121 GGCAATTAAGTCCAGTAACATTTACAGCTTATATTTGTAATCTTCTCCCTGGATATAGA 3180  
Db 3233 GGCAATTAAGTCCAGTAACATTTACAGCTTATATTTGTAATCTTCTCCCTGGATATAGA 3292  
QY 3181 AAGTATCAGCTTAACATTTGATGTCAGAAAGTCTATCAATTTTGGAGTGTGAATTCAGT 3240  
Db 3293 AAGTATCAGCTTAACATTTGATGTCAGAAAGTCTATCAATTTTGGAGTGTGAATTCAGT 3352  
QY 3241 AGAGAAATTTCAAGCAATTAATCTAGCTTACCTTATTAATCTTATGATGTCATGATGGG 3300  
Db 3353 AGAGAAATTTCAAGCAATTAATCTAGCTTACCTTATTAATCTTATGATGTCATGATGGG 3412  
QY 3301 AGTCTTAAACCGAAGAAAGCTTTGAATATGCTGAATCTTGGAGACAGAAAGAAAGTGGC 3360  
Db 3413 AGTCTTAAACCGAAGAAAGCTTTGAATATGCTGAATCTTGGAGACAGAAAGAAAGTGGC 3472  
QY 3361 ATGCAATTTCTGGGTGATCAGAGTCCAAACTTTCGATCTCCGAGCAGCAGCTCCCTG 3420  
Db 3473 ATGCAATTTCTGGGTGATCAGAGTCCAAACTTTCGATCTCCGAGCAGCAGCTCCCTG 3532  
QY 3421 GATATGAAAGTTCAGGCTTAATGCACTGCTCTCACTTCTTAAATTTCAAGCTTCTGAG 3480  
Db 3533 GATATGAAAGTTCAGGCTTAATGCACTGCTCTCACTTCTTAAATTTCAAGCTTCTGAG 3592  
QY 3481 GGAATCCCAATTAATGAGTGGCTTAAGCAGGCAAGAAATAGCTTGGTGGTTTTGATCT 3540  
Db 3593 GGAATCCCAATTAATGAGTGGCTTAAGCAGGCAAGAAATAGCTTGGTGGTTTTGATCT 3652  
QY 3541 ACTCAGATACACTGAGCTTTAAAGGCTCTGCTGAATTTGAGGCTTCAAGCTTCAATGATA 3600  
Db 3653 ACTCAGATACACTGAGCTTTAAAGGCTCTGCTGAATTTGAGGCTTCAAGCTTCAATGATA 3712  
QY 3601 GAAAGCAAAATATCCAAAGTGAACGCTGACGAGGCTTACCAAGTCTCT----- 3651  
Db 3713 GAAAGCAAAATATCCAAAGTGAACGCTGACGAGGCTTACCAAGTCTCTGTAAGTTT 3772  
QY 3652 -----CTTGTGTGTGTAACGCCA 3669  
Db 3773 CTGATTGACACACAACCGCTTACTCTTCAAGCAGCAGAGCTTCTGTGTGTAACGCCA 3832  
QY 3670 ATGGCAGTTAATTTTCCGCAAAATGTTTGAATTTGCTAATTTGTCAGCTCAATGTTGTA 3729  
Db 3833 ATGGCAGTTAATTTTCCGCAAAATGTTTGAATTTGCTAATTTGTCAGCTCAATGTTGTA 3892  
QY 3730 TATATGTAAGGCTTCTGGGTCTCTAGAAAGCAAGATCTATCCAAATCAAGAGCC 3789

```

|||||
Db 3893 TATAAGTAAAGGCTTCTGCTCTTCTTAAAGCGAAGATCTATCAAAATCAAGAGGCC 3952
|||
Qy 3790 TTTGATTTAGATGCTGCTGTAAGAAATTAAGATGATCTCAATCATGNGATTTGAAT 3849
|||
Db 3953 TTTGATTTAGATGCTGCTGTAAGAAATTAAGATGATCTCAATCATGNGATTTGAAT 4012
|||
Qy 3850 GTGTGTACAAAGCTTTTGGGCGCGGTAGAGAGTGGAGCTTTTATGGAAGTTAACTTA 3909
|||
Db 4013 GTGTGTACAAAGCTTTTGGGCGCGGTAGAGAGTGGAGCTTTTATGGAAGTTAACTTA 4072
|||
Qy 3910 TTTAGTGGCTTTATGCTGCTTCAAGAGCAATTTCTCTAGCGAGACAGTGAAGAAATG 3969
|||
Db 4073 TTTAGTGGCTTTATGCTGCTTCAAGAGCAATTTCTCTAGCGAGACAGTGAAGAAATG 4132
|||
Qy 3970 GAATATGATCATGGAAGAAAGCTCAATTTAGATTTCTGTAATGAACCAAGTTTGT 4029
|||
Db 4133 GAATATGATCATGGAAGAAAGCTCAATTTAGATTTCTGTAATGAACCAAGTTTGT 4192
|||
Qy 4030 GTTAATATTTCTGCTGTGAGAAACTTTAAAGTTTCAATACCAAGATGCTTCAAGTCTC 4089
|||
Db 4193 GTTAATATTTCTGCTGTGAGAAACTTTAAAGTTTCAATACCAAGATGCTTCAAGTCTC 4252
|||
Qy 4090 ATATGATGATTTACTATGAGCCAGAGAGACAGCGGTGAGAAATTTCAACTCTGAAGTGAAG 4149
|||
Db 4253 ATATGATGATTTACTATGAGCCAGAGAGACAGCGGTGAGAAATTTCAACTCTGAAGTGAAG 4312
|||
Qy 4150 CTGTCTCTGCTGAGACCTTTGAGTATGTCAGAGGCGTGCCTGCTGTGAGATGAGT 4209
|||
Db 4313 CTGTCTCTGCTGAGACCTTTGAGTATGTCAGAGGCGTGCCTGCTGTGAGATGAGT 4372
|||
Qy 4210 TCAGGCTCCCATCATCATCTTCACTATTTTATTTCTGATTTCAAGCTTCTGATCTT 4269
|||
Db 4373 TCAGGCTCCCATCATCATCTTCACTATTTTATTTCTGATTTCAAGCTTCTGATCTT 4432
|||
Qy 4270 ATGGAACCTTGGCTGTGA 4287
|||
Db 4433 ATGGAACCTTGGCTGTGA 4450
|||

RESULT 4
AX534936 4761 bp DNA linear PAT 22-NOV-2002
LOCUS Sequence 3 from Patent WO02070696.
DEFINITION AX534936
ACCESSION AX534936
VERSION AX534936.1 GI:25261478
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 Schuh, A. and Sutherland, R.D.
Cdi09 nucleic acid molecules polypeptides and methods of use
Patent: WO 02070696-A3 12-SEP-2002;
Schuh, Andre (CA) ; Sutherland, Robert D. (CA)
LOCATION/Qualifiers
1. 4761
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
113..4450
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD57232.1"
/db_xref="GI:25261479"
/translation="MQGPPLTAALHLLCVTAALAAVAPGRFLVTAAGIIRPGANVTI
GVLHCHPSOVTVAQKELKTAENLTVSLAAGVEKESFPTLTPSLPLNSADEIT
ELRVGRDDELFSNSTRLSPETKRSISVFIOTDKALVYKROKURVITLPSDEKPY
KTSNLILIDPKSNLLOQWLSQSDLGAVISKTFPLSSNHLIDGMSI QVAVNDQTYQS
FOVSEYVLEKFEVTLTLPVLCMSNSHGLGTTTAKTYKPKVGDVTLTLELSTWKG
KNITKTFKINSANFSFNDKMNWNSNGLSEYLDLSDSGPEVILTTLESYTGK
SNVSTNVFVKOHDIYIEFPDYTVLKPILNFTATVKTVRADNQLTLERRNNVIT

```

Y  
Res 1

```

VTORNYEYWSGNSGNQKMEAVOKINITYVQSGTFKIEPIEDSSELQKAYFLGS
KSNMAVHSLFKSPKTYVQLTRDENIIVGSPFELVYVGNRLKELSMVNSRODLVA
VKGOMTWESLTPENSWTPKACVLYVYEDGELISDLKXIPVOLVEXKILKWSKY
KABSEKXSLISVTOPDSIVGIYAVVAVKSVLNMASNDIMENYVHELIVNTGYLGS
MFMNSFAYFQCGMLVTLDAUTLDYIDGVYDVAEYARFENBEGHIVLIDHLSLGS
SPHAKHPETWIMLDTNWGSRIYQEEFVYVDSITSVWAEVPIVLEDELGLTTPV
ELOAPPEPFIPLNLPYSVIRGEFALERTINYLKADAEVVIIEKSPKPILMTNE
INAGHOOTLLVPSDEGATVLEPIRPTHLGILPTVPLASPADAOMILVKAEG
IKSXSOSILDLTDNRLOSTLKTLSFSPRPNTVSGSERVOITAGIDVGPISLSTL
IRMPYCGEOMNINPAPVITYLDITTKKQJLTDNLKERALSFMROGYRELLYQEDG
SFSAGNDPESGSTWLSAFVLRCEADVPYIDIDONVLRHTYMLKHKOKSBERWD
GRVHSELQGNKSPVTLTAYIVSLLEGRYOPNDIVQESIHFLSEFSRGISDNVT
LALITVALSNGSPKAKALMLTLRAELEGFASQDFWVSESKLSBMOPRSLDLEVA
YALISHFOFOTSGIPIMRWLSROBNLSGGFASQDFVTLKALSEFAALMNTERTN
IOVWTGSSPSPVKPLIDTNRNLLIOTLAEIYVOPMNVINSANGFALCOLANVYN
VKAGSSRRRSRSIONOAEFDDIVAYKENDLHNDVNLVCSFSGRSGMALEVNIL
LSGFVPSBEALISLETAKVYDHKMLYLDSEVNEFQVNIAPVNRVFXSNVODAS
VSIYDYEPFRQAVRSYNSVEVLSKCDLSDVQCRPCEDAGSGSHHSVIFLFCFK
LLYFMEMLWL"
BASE COUNT 1419 a 913 c 1016 g 1413 t
ORIGIN
Query Match 98.5%; Score 4222.8; DB 6; Length 4761;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 4285; Conservative 0; Mismatches 2; Indels 51; Gaps 1;
1 ATGCAAGGCGCAACGCGCTCTGACCGCGCGCAACCTCTGCTGTGTGACCGCGCGCTG 60
113 ATGCAAGGCGCAACGCGCTCTGACCGCGCGCAACCTCTGCTGTGTGACCGCGCGCTG 172
61 GCGGTGCTCCCGGCGCTCTGCTGTGTGTGACAGCGCGAGGATCATGAGCGCGAGAA 120
173 GCGGTGCTCCCGGCGCTCTGCTGTGTGTGACAGCGCGAGGATCATGAGCGCGAGAA 232
121 AATGTACTATTGGGGGTGAGAGCTTCTGGAACAACCTGCTTCAACAGTGAAGTGAAGCG 180
223 AATGTACTATTGGGGGTGAGAGCTTCTGGAACAACCTGCTTCAACAGTGAAGTGAAGCG 292
181 GAGCTGCTCAAGACAGCATCAACCTCACTGCTCTGCTGTGAGAGAGAGAGTCTTT 240
293 GAGCTGCTCAAGACAGCATCAACCTCACTGCTCTGCTGTGAGAGAGAGAGTCTTT 352
241 GAAAAAGGCTCTTTAAGACACTTACTCTTCATCATCACTCTGACAGTGAAGTGAAG 300
353 GAAAAAGGCTCTTTAAGACACTTACTCTTCATCATCACTCTGACAGTGAAGTGAAG 412
301 ATTATGAGCTACGTTGAACCGGAGCTACCGAGATGAGATTTATTTCTTAATAGTACC 360
413 ATTATGAGCTACGTTGAACCGGAGCTACCGAGATGAGATTTATTTCTTAATAGTACC 472
361 CGCTTATCATTTGAGACCAAGAAATATCTGCTTCAATCAACAGACAGAGCTTATAC 420
473 CGCTTATCATTTGAGACCAAGAAATATCTGCTTCAATCAACAGACAGAGCTTATAC 532
421 AAGCCAAAGCAAGAGTGAAGTTGGATGTTGACACTCTCTGAGATTTTAAAGCTTAC 480
533 AAGCCAAAGCAAGAGTGAAGTTGGATGTTGACACTCTCTGAGATTTTAAAGCTTAC 592
481 AAAACCTCTTTAAACATTTCTATTAAGAGACCCCAATCAAAATTTGATCAACAGTGTG 540
593 AAAACCTCTTTAAACATTTCTATTAAGAGACCCCAATCAAAATTTGATCAACAGTGTG 652
541 TCACAACAAAGTGAATCTTGAAGTCAATTTCCAAAATTTTCAAGCTATCTTCCATGCAATA 600
653 TCACAACAAAGTGAATCTTGAAGTCAATTTCCAAAATTTTCAAGCTATCTTCCATGCAATA 712
601 CTGTGTGCTGCTCATTAAGTCAAGTGAATGACGAGATVTTATCAATCATTTAG 660
713 CTGTGTGCTGCTCATTAAGTCAAGTGAATGACGAGATVTTATCAATCATTTAG 772
661 GTTTCAGAAATATGATTAACAAATTTGAAGTGAATTTGACAGACCATTAATATTTCT 720
773 GTTTCAGAAATATGATTAACAAATTTGAAGTGAATTTGACAGACCATTAATATTTCT 832

```

QY 721 ATGAATCTAAGCATTTAAATGTATCAATCAAGCAAGTATATATATGGAAGCCAGTG 780  
DB 833 ATGAATCTAAGCATTTAAATGTATCAATCAAGCAAGTATATATATGGAAGCCAGTG 892  
QY 781 AAAGAGACGTAAAGCCTTACATTTTACCTTTATCTTTTGGGAAAGAGAAAAATTT 840  
DB 893 AAAGAGACGTAAAGCCTTACATTTTACCTTTATCTTTTGGGAAAGAGAAAAATTT 952  
QY 841 ACAAACAATTAAAGATTAATGTATGCAACTCTCTTTTAAATGATGAAGATGAAA 900  
DB 953 ACAAACAATTAAAGATTAATGTATGCAACTCTCTTTTAAATGATGAAGATGAAA 1012  
QY 901 AATGTATGATTTCTTCAATATGACCTTCTGATATCTGATATCTTCTCCCTGACCA 960  
DB 1013 AATGTATGATTTCTTCAATATGACCTTCTGATATCTGATATCTTCTCCCTGACCA 1072  
QY 961 GTAGAAATTTTAAACCAAGTACAGATCACTTACAGATATTTCAAGAAATGTAGCACT 1020  
DB 1073 GTAGAAATTTTAAACCAAGTACAGATCACTTACAGATATTTCAAGAAATGTAGCACT 1132  
QY 1021 AATGTATCTTCAAGCAATGATTAATCATGATGATTTTGTATTAATTAATTAATTAAT 1080  
DB 1133 AATGTATCTTCAAGCAATGATTAATCATGATGATTTTGTATTAATTAATTAATTAAT 1192  
QY 1081 AAGCATCTCTCAACTTCAAGCACTGTGAAAGTAACTGCTGATGATGCAACCACTG 1140  
DB 1193 AAGCATCTCTCAACTTCAAGCACTGTGAAAGTAACTGCTGATGATGCAACCACTG 1252  
QY 1141 ACTCTTGAAGAAAGAAATTAATGTATCTTAAACAGTACACAGAAATTAATTAAT 1200  
DB 1253 ACTCTTGAAGAAAGAAATTAATGTATCTTAAACAGTACACAGAAATTAATTAAT 1312  
QY 1201 TACTGAGCGGATCTAAACAGTGAATTCAGAAATGGAAGCTGTCAGAAATTAATTAAT 1260  
DB 1313 TACTGAGCGGATCTAAACAGTGAATTCAGAAATGGAAGCTGTCAGAAATTAATTAAT 1372  
QY 1261 ACTGTCCCCCAAGGTGAACTTTTAAATGATTAATCCCAATCTGAGAGATTCAGAGT 1320  
DB 1373 ACTGTCCCCCAAGGTGAACTTTTAAATGATTAATCCCAATCTGAGAGATTCAGAGT 1432  
QY 1321 CTACAGTTGAAGGCTTATTTCTTGTGTATTAAGTAAAGTACATGATGATGATGAT 1380  
DB 1433 CTACAGTTGAAGGCTTATTTCTTGTGTATTAAGTAAAGTACATGATGATGATGAT 1492  
QY 1381 AAGTCTCTAGTAAAGATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1440  
DB 1493 AAGTCTCTAGTAAAGATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1552  
QY 1441 TCGCCTTTTGAAGTGTGTGTATGTGCAAAAGATTTGAAGATTAAGTATATGAT 1500  
DB 1553 TCGCCTTTTGAAGTGTGTGTATGTGCAAAAGATTTGAAGATTAAGTATATGAT 1612  
QY 1501 GTATTCAGAGGACAGTGTGTGTGTATGTGCAAAAGATTTGAAGATTAAGTATATGAT 1560  
DB 1613 GTATTCAGAGGACAGTGTGTGTGTATGTGCAAAAGATTTGAAGATTAAGTATATGAT 1672  
QY 1561 CCAAAAATTTCTTGAATCTCCAAAAGCTGTGTATTTGTATTAATTTGAAGATGATG 1620  
DB 1673 CCAAAAATTTCTTGAATCTCCAAAAGCTGTGTATTTGTATTAATTTGAAGATGATG 1732  
QY 1621 GAAATTTAAAGTATGTTCTTAAATTTCTGTGATGATGTTTAAATTAAGTATGAT 1680  
DB 1733 GAAATTTAAAGTATGTTCTTAAATTTCTGTGATGATGTTTAAATTAAGTATGAT 1792  
QY 1681 CTATATTTGAAGTAAAGTGAAGCTGAACCATCTGAGAAAGTCTCTTGAAGATCTGTG 1740  
DB 1793 CTATATTTGAAGTAAAGTGAAGCTGAACCATCTGAGAAAGTCTCTTGAAGATCTGTG 1852  
QY 1741 ACAAGGCTGATCTCATATGTTGGATTTGTAGCTTTGACAAATTTGTGAATCTGATGAT 1800  
DB 1853 ACAAGGCTGATCTCATATGTTGGATTTGTAGCTTTGACAAATTTGTGAATCTGATGAT 1912

QY 1801 GCCTTAATGATATTAATCAATGAAATGTGTCCATGATGATGATGATGATGATGATGAT 1860  
DB 1913 GCCTTAATGATATTAATCAATGAAATGTGTCCATGATGATGATGATGATGATGATGAT 1972  
QY 1861 TATTAATTTGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920  
DB 1973 TATTAATTTGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2032  
QY 1921 TTGACAGATGCAAACTCAAGAAAGATTAATGATGATGATGATGATGATGATGATGAT 1980  
DB 2033 TTGACAGATGCAAACTCAAGAAAGATTAATGATGATGATGATGATGATGATGATGAT 2092  
QY 1981 GCTGAAGTATTAAGAGAAATGAAGCATATTTGTATGATATTAATGATGATGATGAT 2040  
DB 2093 GCTGAAGTATTAAGAGAAATGAAGCATATTTGTATGATATTAATGATGATGATGAT 2152  
QY 2041 GTTAGAGTCCCAATGTCGAAAGATTTTCCAGAGCTGATGATGATGATGATGATGATGAT 2100  
DB 2153 GTTAGAGTCCCAATGTCGAAAGATTTTCCAGAGCTGATGATGATGATGATGATGATGAT 2212  
QY 2101 ATGGGTTACAGATTTTACCAAGATTTGAAGTAACTGATCTGATCTATCACTTCTTGG 2160  
DB 2213 ATGGGTTACAGATTTTACCAAGATTTGAAGTAACTGATCTGATCTATCACTTCTTGG 2272  
QY 2161 GTGGCTACGCTTTTGTATCTCTGAGACCTGGGCTTGTGATTAACATCACTCACTG 2220  
DB 2273 GTGGCTACGCTTTTGTATCTCTGAGACCTGGGCTTGTGATTAACATCACTCACTG 2332  
QY 2221 GAGCTCAAGCTTCCCAATTTTCAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGA 2280  
DB 2333 GAGCTCAAGCTTCCCAATTTTCAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGA 2392  
QY 2281 GGTGAAGATTTTCTTGAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2340  
DB 2393 GGTGAAGATTTTCTTGAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2452  
QY 2341 AAGTATATCTTGAAGAAAGTGAACAAATTTGATTTCTAATGATCTTCAAGTAAATTAAT 2400  
DB 2453 AAGTATATCTTGAAGAAAGTGAACAAATTTGATTTCTAATGATCTTCAATGAAATTAAT 2512  
QY 2401 GCCACAGGCTCACAGAGACCTTCTGTGTGTTCCAGTGAAGATTTGGGCAACTGTTCTTTT 2460  
DB 2513 GCCACAGGCTCACAGAGACCTTCTGTGTGTTCCAGTGAAGATTTGGGCAACTGTTCTTTT 2572  
QY 2461 CCCATCAGGCTCACATCTGGGAGAAATTCATATCAAGTCAAGCTCTTTCAACCACCT 2520  
DB 2573 CCCATCAGGCTCACATCTGGGAGAAATTCATATCAAGTCAAGCTCTTTCAACCACCT 2632  
QY 2521 GCTTCTGATGCTGTCAACCATGATTTTGAAGGCTGAAGGATTAAGAAATCAATAT 2580  
DB 2633 GCTTCTGATGCTGTCAACCATGATTTTGAAGGCTGAAGGATTAAGAAATCAATAT 2692  
QY 2581 TCAATATCATCTTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2640  
DB 2693 TCAATATCATCTTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2752  
QY 2641 AGTTTCAATTTCTCTTAATTAAGTACAGTACAGTACAGTACAGTACAGTACAGTACAT 2700  
DB 2753 AGTTTCAATTTCTCTTAATTAAGTACAGTACAGTACAGTACAGTACAGTACAGTACAT 2812  
QY 2701 GGAAGATTTCTTGTGCTTCCATCAATGATGATGATGATGATGATGATGATGATGATG 2760  
DB 2813 GGAAGATTTCTTGTGCTTCCATCAATGATGATGATGATGATGATGATGATGATGATG 2872  
QY 2761 TGTGTGAACAGAACTGATTAATTTTGTCTCCAAATTTTGAATTTTGAATTTTGAATTT 2820  
DB 2873 TGTGTGAACAGAACTGATTAATTTTGTCTCCAAATTTTGAATTTTGAATTTTGAATTT 2932  
QY 2821 AAAAAGAAACAACTGACAGTATTTGAAGAAAGCTCTTCAATTTTGAAGCAAGT 2880  
DB 2933 AAAAAGAAACAACTGACAGTATTTGAAGAAAGCTCTTCAATTTTGAAGCAAGT 2992  
QY 2881 TACAGAGAACTTCTATACAGAGGAAAGATGCTCTTTCAGTCTTTTGGAAATTAAT 2940

```

Db      2993  TACCAAGAGAACTTCTATCAGAGGAGAGATGCGCTCTTTCAGTCCTTTGGGAATTAAT
Qy      2941  GACCCCTTGGGAGACCTTGCTGAGCTTTGTTTAAGATGTTCCCTGAAGCCGAT
Db      3053  GACCCCTTGGGAGACCTTGCTGAGCTTTGTTTAAGATGTTCCCTGAAGCCGAT
Qy      3001  CCTTACATAGATATTGATCAGAAATGTTTACACAGAACATACCTTGCTTAAGAGCAT
Db      3113  CCTTACATAGATATTGATCAGAAATGTTTACACAGAACATACCTTGCTTAAGAGCAT
Qy      3061  CAGAAATCCAGCGGTGAATTTGGATCCAGAGAGGTGATTCATATGTAAGCTTCAAGGT
Db      3173  CAGAAATCCAGCGGTGAATTTGGATCCAGAGAGGTGATTCATATGTAAGCTTCAAGGT
Qy      3121  GGCATTAAGATCAGTAAACCTTACAGGCTTATTTGATCTCTCCGCGGATTTAA
Db      3233  GGCATTAAGATCAGTAAACCTTACAGGCTTATTTGATCTCTCCGCGGATTTAA
Qy      3181  AAGTATCAGCTTAACTTGAATGTCAGAGAGCTTATCATTGTTTGGAGCTGATTCAGT
Db      3293  AAGTATCAGCTTAACTTGAATGTCAGAGAGCTTATCATTGTTTGGAGCTGATTCAGT
Qy      3241  AAGAGATTTTCAACAATTATCTTACGCTTATTAATCTTATGATTCATCAGTGGG
Db      3353  AAGAGATTTTCAACAATTATCTTACGCTTATTAATCTTATGATTCATCAGTGGG
Qy      3301  AGTCCCAAAGCGAAGAGCTTGAATATGCTGACCTTGAAGAGCAACAAAGAGTGGC
Db      3413  AGTCCCAAAGCGAAGAGCTTGAATATGCTGACCTTGAAGAGCAACAAAGAGTGGC
Qy      3361  ATGCAATTCGGGTGTCATCAGAGTCCAAACTTTCGACTCCGCGAGCAAGCTCCG
Db      3473  ATGCAATTCGGGTGTCATCAGAGTCCAAACTTTCGACTCCGCGAGCAAGCTCCG
Qy      3421  GATATTAAGTTGACGCTTATGACCTGCTCTCACTTCTTCAATTTGACAGCTTGA
Db      3533  GATATTAAGTTGACGCTTATGACCTGCTCTCACTTCTTCAATTTGACAGCTTGA
Qy      3481  GGAATTCCTCAATTAATGAGTGGCTTAAGAGAGCAAGAAAGATGTTGGTGGTTCAT
Db      3593  GGAATTCCTCAATTAATGAGTGGCTTAAGAGAGCAAGAAAGATGTTGGTGGTTCAT
Qy      3541  ACTCAGATACCACTGCTGCTTAAAGCTGCTGCTGATTTGACAGCCCTTAATGAATACA
Db      3653  ACTCAGATACCACTGCTGCTTAAAGCTGCTGCTGATTTGACAGCCCTTAATGAATACA
Qy      3601  GAAAGACAATATTCAGATGACCGTGACGGGGCTTACCTCAACCAAGTCT-----
Db      3713  GAAAGACAATATTCAGATGACCGTGACGGGGCTTACCTCAACCAAGTCT-----
Qy      3652  -----CTTGTCTGTGTGACAGCA
Db      3773  CTGATTTGACACACAAACCGCTTACTCTTCAACAGACAGAGCTGCTGTGTGACAGCA
Qy      3670  ATGCGAGTTAATATTTCCGCAATATGTTTGAATTTGCTATTTGCTGAGCTCAATGTTGA
Db      3833  ATGCGAGTTAATATTTCCGCAATATGTTTGAATTTGCTATTTGCTGAGCTCAATGTTGA
Qy      3730  TATATGTAAGAGCTTCTGGGCTTCTTCAAGAGCAAGATCTTCAATCAAGAGAGCC
Db      3893  TATATGTAAGAGCTTCTGGGCTTCTTCAAGAGCAAGATCTTCAATCAAGAGAGCC
Qy      3790  TTTGATTTGATGTTGCTGTTAAAGAAAATAAGATGATCTCAATCATGTGATTTGAAT
Db      3953  TTTGATTTGATGTTGCTGTTAAAGAAAATAAGATGATCTCAATCATGTGATTTGAAT
Qy      3850  GTGTGTACAGCTTTTCCGGCCCGGGTAGAGAGTGATGCTCTTATGAGATTTAACCTA
Db      4013  GTGTGTACAGCTTTTCCGGCCCGGGTAGAGAGTGATGCTCTTATGAGATTTAACCTA
Qy      3910  TTAAGTGCTTTATGTTGCTTCAAGAGCAATTTCTCTGAGGAGACAGTGAAGAAAGTG

```

Db 4073 TTAAGTGCTTTATGTTGCTTCAAGAGCAATTTCTTGAAGAGACAGTGAAGAAAGTG 4132

Qy 3970 GAATATGATCATGGAACCTCAACCTTAATTTAGATTCCTGTAATGAACCAAGTTTGT 4029

Db 4133 GAATATGATCATGGAACCTCAACCTTAATTTAGATTCCTGTAATGAACCAAGTTTGT 4192

Qy 4030 GTTAATATTTCTGCTGTGAGAACTTTAAAGTTTCAATTAACCAAGATGCTTCAAGTTC 4089

Db 4193 GTTAATATTTCTGCTGTGAGAACTTTAAAGTTTCAATTAACCAAGATGCTTCAAGTTC 4252

Qy 4090 ATATGATTTACTATGAGCCAAAGAGACAGCGCGTGAGAGTTTACATCTGAGAGTGAAG 4149

Db 4253 ATATGATTTACTATGAGCCAAAGAGACAGCGCGTGAGAGTTTACATCTGAGAGTGAAG 4312

Qy 4150 CTGCTCCTGTCGACCTTTGAGATGATCCAGAGCGCTGCGCTTGTGAGAGATGAGACT 4209

Db 4313 CTGCTCCTGTCGACCTTTGAGATGATCCAGAGCGCTGCGCTTGTGAGAGATGAGACT 4372

Qy 4210 TCAGGCTCCATCATCAGCTTTCAGTCAATTTTATTTTCTGTTTCAAGCTTCTGACTTT 4269

Db 4373 TCAGGCTCCATCATCAGCTTTCAGTCAATTTTATTTTCTGTTTCAAGCTTCTGACTTT 4432

Qy 4270 ATGGAACCTTGGCTGTGA 4287

Db 4433 ATGGAACCTTGGCTGTGA 4450

RESULT 5  
AX534940 5895 bp DNA linear PAT 22-NOV-2002  
LOCUS AX534940  
DEFINITION Sequence 7 from Patent WO02070696.  
ACCESSION AX534940  
VERSION AX534940.1 GI:25261486  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1  
Schuh, A. and Sutherland, R.D.  
Cd109 nucleic acid molecules polypeptides and methods of use  
Patent: WO 02070696-A 7 12-SEP-2002;  
Schuh, Andre (CA); Sutherland, Robert D. (CA)

FEATURES  
source location/Qualifiers  
1..5895  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
113..4450  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD57234.1"  
/db\_xref="gi:25261487"  
/translation="MCGPPLITLTAHLICVCTAALAVAPGPFVLTAPGIIIRGGNVTL  
GVELLEHCPQVTVKABELTKTASNLIVTSVAEGVFEKSGFKTLTLPISNADETV  
ELAVTGRQDILLNSNSTRLSPEKRLSVPTQTKALYKPKQYKFRVTLPSDPKY  
KTSNLITLKDKSNLIQWLSQSDLDGYSKTFQLSHPILIGDSIQVQVDQTYQS  
FOVSEYVLPKEVTLQTPLYCSNMSKHLGITLAKYKQKPKVDVTLTFLPSWGK  
KNKITKFKINGSANFSEFNDEKKNVMSNGLESDYDLSGPGVEILITVTESVYGI  
SRNVSTNVEFQHHYITIEFPDYTVLKSLNFTATVKTADGNOLTLERRNNVIT  
VTORNYTFMGSNGNOKMEAVOKINTVTVDSGTFKIEFPILEDSSRLQKAYLGS  
KSSNAHSLFSPSKTYIQLKTRBNIKVSGPFLVYSGNRLKELSTMYVSROQVA  
VGRQSTVSLTPENSWTPKACVITYIEDGEITISDLKIPVQLVFNKRLKLTWSKV  
KAEPSEKSLRISVQDPSIVGIVAVDSVNMNANSNDITMENVHELELNTGYLIG  
MENMSFAVQECGMVLVDANLTQYIDGVNDNAEYERFMEBNGHILVDJHDSIGS  
SPHAKHPPEFTWIMLDTNMGSRIVQEFVTVDSITSVATGPFVISEDLAGLTTTPV  
ELQAFQPEPFLNLPYGVIRGEERALETTFNVLKDAETEVILIEKSDKPIIMTNS  
IMATGHOOTLLVPSDEGATVLPTRPHLGIPTPTTALSTPAADATQMLIVABGI  
EKYSQSTLLDLNRLQSTLKTISFSFPEPTVGSERVQITAGDVLSGINSASL  
IRMPGVCBQNMNINPAPNITLIDVLTAKKQVLTIDQNVLHRTYWLKQKNGEFTWD  
SRVHSELOGNKSPTVLTAVITVSLGKYCQPNIDQESIHFLSEFSRGI SDNYT  
LALITVLSVSGPKAKALNMLTWBAEQEGMQFWVSSSEKLSDSMQPRDLIEVAA

Same as  
Result 1

YALLSHFLQFQSEGIPIMRWLSHORSNLSGGFASDTDTVALKALSEFALMINTERN  
 IOVTVTPSSSPVKPLIDITHNRLLOTAELAVVOPMAVINISAGFEFALCOLAVVN  
 VKASRRRSRRIIONEAFDLVDVAVKENKDINHNDVNSFSGPERSGMALEVL  
 LSGMVBSEATLSLSETKVYEDHKLNLYDSVNEPOFCNIPAVNPFVSNTPDSS  
 VSIYDYEPRRQAVRSYNSBYKLSCDLCSVQCRFCBDSAGSHHSVIFPFCRK  
 LHYMEML"

BASE COUNT 1708 a 1174 c 1241 g 1772 t  
 ORIGIN

Query Match 98.5%; Score 4222.8; DB 6; Length 5895;  
 Best Local Similarity 98.8%; Pred. No. 0;  
 Matches 4285; Conservative 0; Mismatches 2; Indels 51; Gaps 1;

Qy 1 ATGAGAGGCCACCGCTCTGACCGCGCCACCTCTCTGCGTGTGACCGCGCGCTG 60  
 Db 113 ATGAGAGGCCACCGCTCTGACCGCGCCACCTCTCTGCGTGTGACCGCGCGCTG 172  
 Qy 61 GCCGTGCTCCCGGCTCTGCTGCTGACAGCCCGGAGATCATAGGCGCGAGGA 120  
 Db 173 GCCGTGCTCCCGGCTCTGCTGCTGACAGCCCGGAGATCATAGGCGCGAGGA 232  
 Qy 121 AATGACTATTGGGAGGAGCTCTGGAACCTGCGCTTCAAGGTGACTGTGAAGGCG 180  
 Db 233 AATGACTATTGGGAGGAGCTCTGGAACCTGCGCTTCAAGGTGACTGTGAAGGCG 292  
 Qy 181 GAGCTGCTCAAGACAGATCAAACTCATCTCTCTGCTGGAAGAGAGAGTCTTT 240  
 Db 293 GAGCTGCTCAAGACAGATCAAACTCATCTCTCTGCTGGAAGAGAGAGTCTTT 352  
 Qy 241 GAAAAAGGCTCTTTTAAAGACCTTACTCTTCTCATCACTACTCTGAAAGTGAGATGAG 300  
 Db 353 GAAAAAGGCTCTTTTAAAGACCTTACTCTTCTCATCACTACTCTGAAAGTGAGATGAG 412  
 Qy 301 ATTATAGCTACGTGTAACCGGACGTACCGAGATGAGATTTATCTCTTAATAGTACC 360  
 Db 413 ATTATAGCTACGTGTAACCGGACGTACCGAGATGAGATTTATCTCTTAATAGTACC 472  
 Qy 361 CGCTTATCATTTGAGACCAAGAGATATCTGTCTTCAATTCAAGACAAAGGCTTTATAC 420  
 Db 473 CGCTTATCATTTGAGACCAAGAGATATCTGTCTTCAATTCAAGACAAAGGCTTTATAC 532  
 Qy 421 AAGCCAAAGCAAGAGTGAAGTTTCGCAATGTTTACACTCTTCTGCAATTTTAAAGCTTTAC 480  
 Db 533 AAGCCAAAGCAAGAGTGAAGTTTCGCAATGTTTACACTCTTCTGCAATTTTAAAGCTTTAC 592  
 Qy 481 AAAACCTCTTAAACATTTCTCATTAAGGACCCCAATCAAAATTGATCCAAAGTGTG 540  
 Db 593 AAAACCTCTTAAACATTTCTCATTAAGGACCCCAATCAAAATTGATCCAAAGTGTG 652  
 Qy 541 TCACAAAGAGTGAATCTTGAGTCAATTTCCAAACTTTTCACTATCTCCATCCATA 600  
 Db 653 TCACAAAGAGTGAATCTTGAGTCAATTTCCAAACTTTTCACTATCTCCATCCATA 712  
 Qy 601 CTGTGTGACTGTGTATTCAGAGTTCAAGTGAATGACAGACATATATCAATCAATTTTCAAG 660  
 Db 713 CTGTGTGACTGTGTATTCAGAGTTCAAGTGAATGACAGACATATATCAATCAATTTTCAAG 772  
 Qy 661 GTTTCAGAAATGATTAACCAAAATTGAGAGTGAATTTGAGAGACATATATATTTCTT 720  
 Db 773 GTTTCAGAAATGATTAACCAAAATTGAGAGTGAATTTGAGAGACATATATATTTCTT 832  
 Qy 721 ATGAATTTCAAGATTTAAATGTATCAATCAAGCAAGTATACATATGAGGAGCCAGTG 780  
 Db 833 ATGAATTTCAAGATTTAAATGTATCAATCAAGCAAGTATACATATGAGGAGCCAGTG 892  
 Qy 781 AAAGAGAGCTGTAAGCTTACATTTTATCTTTATCTTTTGGGAGAAAGAAATATAT 840  
 Db 893 AAAGAGAGCTGTAAGCTTACATTTTATCTTTTGGGAGAAAGAAATATAT 952  
 Qy 841 ACAAACCATTTAAGATTAATGATCTGCAAACTTCTCTTTTATGATGAAGAGATGAAA 900  
 Db 953 ACAAACCATTTAAGATTAATGATCTGCAAACTTCTCTTTTATGATGAAGAGATGAAA 1012

Qy 901 AATGTATGAGATCTTCAATGAGACTTCTGAATACCTGATCTATCTTCCCTGAGCA 960  
 Db 1013 AATGTATGAGATCTTCAATGAGACTTCTGAATACCTGATCTATCTTCCCTGAGCA 1072  
 Qy 961 GTGAAATTTTAAACACAGTGAAGATGATTAACAGTATTTTCAAGAAATGTAACACT 1020  
 Db 1073 GTGAAATTTTAAACACAGTGAAGATGATTAACAGTATTTTCAAGAAATGTAACACT 1132  
 Qy 1021 AATGTGTTCTTCAAGAACATGATTTACATCAATGAGTTTGTGATTAATCACTATCTT 1080  
 Db 1133 AATGTGTTCTTCAAGAACATGATTTACATCAATGAGTTTGTGATTAATCACTATCTT 1192  
 Qy 1081 AAGCCATCTCTCAACTTCAAGCAAGCACTGTGAAGTAACTGTGCTGAGCAACAATG 1140  
 Db 1193 AAGCCATCTCTCAACTTCAAGCAAGCACTGTGAAGTAACTGTGCTGAGCAACAATG 1252  
 Qy 1141 ACTCTTGAAGAAAGAAATATATGATGATTAACAGTGAACAGAGAAACTTACTGAG 1200  
 Db 1253 ACTCTTGAAGAAAGAAATATATGATGATTAACAGTGAACAGAGAAACTTACTGAG 1312  
 Qy 1201 TACTGAGCGGATCTAAAGTGAATCAAGAAATGGAAGCTGTGAGAAATTAATTAAT 1260  
 Db 1313 TACTGAGCGGATCTAAAGTGAATCAAGAAATGGAAGCTGTGAGAAATTAATTAAT 1372  
 Qy 1261 ACTGCCCCCAAGTGAATCTTTAAGATTAATCCAAATCTGAGAGATTCAGTGAG 1320  
 Db 1373 ACTGCCCCCAAGTGAATCTTTAAGATTAATCCAAATCTGAGAGATTCAGTGAG 1432  
 Qy 1321 CTACAGTTGAAGGCTATTTCTTGTGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1380  
 Db 1433 CTACAGTTGAAGGCTATTTCTTGTGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1492  
 Qy 1381 AAGTCTCTAGTGAAGCATATCACTCAACTCAAAAGAGATGAAGAAATTAAGTGGGA 1440  
 Db 1493 AAGTCTCTAGTGAAGCATATCACTCAACTCAAAAGAGATGAAGAAATTAAGTGGGA 1552  
 Qy 1441 TCGCCTTTTGAAGTGTGTGATGAGCAACAAAGATGAAGGATTAAGTGAATGTA 1500  
 Db 1553 TCGCCTTTTGAAGTGTGTGATGAGCAACAAAGATGAAGGATTAAGTGAATGTA 1612  
 Qy 1501 GTATCCAGGGGACAGTTGGTGTGATGAGAAACAAATTTCAACAATGTTCTTTTAACA 1560  
 Db 1613 GTATCCAGGGGACAGTTGGTGTGATGAGAAACAAATTTCAACAATGTTCTTTTAACA 1672  
 Qy 1561 CGAAGAAATCTTGAAGCTCCAAAGCTGTGATTTGTGATTAATGTAAGTGAATGAG 1620  
 Db 1673 CGAAGAAATCTTGAAGCTCCAAAGCTGTGATTTGTGATTAATGTAAGTGAATGAG 1732  
 Qy 1621 GAATTTATAGTATGTTCTTAAAAATTCCTGTTCAGCTTGTTTTAAAAATAGATTAAG 1680  
 Db 1733 GAATTTATAGTATGTTCTTAAAAATTCCTGTTCAGCTTGTTTTAAAAATAGATTAAG 1792  
 Qy 1681 CTATATTGAGATTAAGTGAAGTGAACCACTGAGAAAGTCTCTTTAGAGTCTGTG 1740  
 Db 1793 CTATATTGAGATTAAGTGAAGTGAACCACTGAGAAAGTCTCTTTAGAGTCTGTG 1852  
 Qy 1741 ACAAGCTGAGCTCCATAGTGGGATTTGAGCTGTGACAAAGTGAATGATGATGAAT 1800  
 Db 1853 ACAAGCTGAGCTCCATAGTGGGATTTGAGCTGTGACAAAGTGAATGATGATGAAT 1912  
 Qy 1801 GCCTCTAATGATTAATCAATGGAATAATGTGTCCATAGTGAACCTTTATTAACAGAGA 1860  
 Db 1913 GCCTCTAATGATTAATCAATGGAATAATGTGTCCATAGTGAACCTTTATTAACAGAGA 1972  
 Qy 1861 TATTTATTAAGCATGTTCAATGAATTTCTTTTCAAGTCTTTCAAGAAATGAGACTGTGGTA 1920  
 Db 1973 TATTTATTAAGCATGTTCAATGAATTTCTTTTCAAGTCTTTCAAGAAATGAGACTGTGGTA 2032  
 Qy 1921 TTGACAGATGCAAACTCAAGAGATTAATGATGAGTGTGTTATGACAAATGCAAAATAT 1980  
 Db 2033 TTGACAGATGCAAACTCAAGAGATTAATGATGAGTGTGTTATGACAAATGCAAAATAT 2092  
 Qy 1981 GCTGAGAGGTTTATGAGAGAAATGAGAGCATATTTGATGATTAATGATGATCTTTCTTTG 2040



Db 2093 GCTGAGAGGTTATGAGAGGAAATGAGACATATTTGATGATATTCATGACTTTCTTTG 2152  
Qy 2041 GGTAGAGTCCACATGTCGAAAGATTTTCCAGAGCTTGATTTGGCTAGACACCAAC 2100  
Db 2153 GGTAGAGTCCACATGTCGAAAGATTTTCCAGAGCTTGATTTGGCTAGACACCAAC 2212  
Qy 2101 ATGGGTTACAGATTTTACAGAGATTTGAAATGTAACCTGATCTATCACTTCTTG 2160  
Db 2213 ATGGGTTACAGATTTTACAGAGATTTGAAATGTAACCTGATCTATCACTTCTTG 2272  
Qy 2161 GTGGCTACGTTTGTGATCTCTGAGACCTGGGCTTGAGACTAACCTACTCCAGTG 2220  
Db 2273 GTGGCTACGTTTGTGATCTCTGAGACCTGGGCTTGAGACTAACCTACTCCAGTG 2332  
Qy 2221 GAGCTCAAGCCTTCCACCATTTTCACTTTTGAATCTTCCACTCTGTTATCAGA 2280  
Db 2333 GAGCTCAAGCCTTCCACCATTTTCACTTTTGAATCTTCCACTCTGTTATCAGA 2392  
Qy 2281 GGTGAAGAAATTTGCTTGGAAATTAATAATTAATTTGAAAGATGCCAGTGAGTT 2340  
Db 2393 GGTGAAGAAATTTGCTTGGAAATTAATAATTAATTTGAAAGATGCCAGTGAGTT 2452  
Qy 2341 AAGGTATCATTTGAGAAAGATGCAAAATTTGATTTCTATGACTTCAAGTGAATTAAT 2400  
Db 2453 AAGGTATCATTTGAGAAAGATGCAAAATTTGATTTCTATGACTTCAATGAATTAAT 2512  
Qy 2401 GCCACAGGCCACAGAGACCCCTTGGTTCCAGTAGAGATGGGCCACTGTTCTTTT 2460  
Db 2513 GCCACAGGCCACAGAGACCCCTTGGTTCCAGTAGAGATGGGCCACTGTTCTTTT 2572  
Qy 2461 CCCATAGGCCCAACACATCTGGAGAAATTTCTATCAAGTCAAGCTCTTTCACCCACT 2520  
Db 2573 CCCATAGGCCCAACACATCTGGAGAAATTTCTATCAAGTCAAGCTCTTTCACCCACT 2632  
Qy 2521 GCTTCTGATGCTGTCAACCAAGATTTTATGTAAGCTGAGAGAAATGAAATCATAT 2580  
Db 2633 GCTTCTGATGCTGTCAACCAAGATTTTATGTAAGCTGAGAGAAATGAAATCATAT 2692  
Qy 2581 TCACATTCATCTTATTTAGACTTGACTGACATAGGCTACAGATCCCTGAAATCTTG 2640  
Db 2693 TCACATTCATCTTATTTAGACTTGACTGACATAGGCTACAGATCCCTGAAATCTTG 2752  
Qy 2641 AGTTTCTCACTTCTCTAATAACAGTACGAGCTGGAGAAAGATTGAGATCACTGCAAT 2700  
Db 2753 AGTTTCTCACTTCTCTAATAACAGTACGAGCTGGAGAAAGATTGAGATCACTGCAAT 2812  
Qy 2701 GAGATGTTCTTGTCTTTCATCAATGAGCTTATGCTCATTTGATGAGTGCCTTAATG 2760  
Db 2813 GAGATGTTCTTGTCTTTCATCAATGAGCTTATGCTCATTTGATGAGTGCCTTAATG 2872  
Qy 2761 TGTGTGAACAGAACTGATTAATTTTGTCCAAATTTTATCACTTTGATATCTGACT 2820  
Db 2873 TGTGTGAACAGAACTGATTAATTTTGTCCAAATTTTATCACTTTGATATCTGACT 2932  
Qy 2821 AAAAAGAAACAACTGACAGATTAATTTGAAAGAAAGCTCTTTCATTTATGAGCAAGT 2880  
Db 2933 AAAAAGAAACAACTGACAGATTAATTTGAAAGAAAGCTCTTTCATTTATGAGCAAGT 2992  
Qy 2881 TACAGAGAGAACTTCTATATCAGAGAGAGATGCTCTTTCAGTCTTTTGGAAATTAAT 2940  
Db 2993 TACAGAGAGAACTTCTATATCAGAGAGAGATGCTCTTTCAGTCTTTTGGAAATTAAT 3052  
Qy 2941 GACCTTCTGGAGACCTTGTTGTGACGTTTGTATTAAGATTTTCTTGAAGCCGAT 3000  
Db 3053 GACCTTCTGGAGACCTTGTTGTGACGTTTGTATTAAGATTTTCTTGAAGCCGAT 3112  
Qy 3001 CCTTACATAGATTTGATCAGATGATGTTTACACAGAAACATACCTTGCTTAAAGCAAT 3060  
Db 3113 CCTTACATAGATTTGATCAGATGATGTTTACACAGAAACATACCTTGCTTAAAGCAAT 3172  
Qy 3061 CAGAAATCCACGAGTGAATTTTGGATTCAGAGAGAGATTCATATGATGACTTCAAGT 3120

Db 3173 CAGAAATCCACGAGTGAATTTTGGATTCAGAGAGATGATTCATATGAGCTTCAAGT 3232  
Qy 3121 GGCATTTAAAGTCAGTAACACTTACAGCTTATATTTGATCTTCTCGGATATAGA 3180  
Db 3233 GGCATTTAAAGTCAGTAACACTTACAGCTTATATTTGATCTTCTCGGATATAGA 3292  
Qy 3181 AAGTATCAGCTTACATTTGATGTCAGAGATCTATCCATTTTGTGAGTCTGAATTCAGT 3240  
Db 3293 AAGTATCAGCTTACATTTGATGTCAGAGATCTATCCATTTTGTGAGTCTGAATTCAGT 3352  
Qy 3241 AGAGAAATTTGAGAAATTAATCTTACCCCTTAATCTATGATTTGATCACTGAGGG 3300  
Db 3353 AGAGAAATTTGAGAAATTAATCTTACCCCTTAATCTTACCTTATGATTTGATCACTGAGGG 3412  
Qy 3301 AGTCTTAAAGCAGAAAGATTTGATATGCTGACTTGAAGAGAGCAACAGAGTGGC 3360  
Db 3413 AGTCTTAAAGCAGAAAGATTTGATATGCTGACTTGAAGAGAGCAACAGAGTGGC 3472  
Qy 3361 ATGCAATTTCTGGGTCTATCAGAGTCCAAACTTTCTGACTCTGGCAGCCAGCTCCCTG 3420  
Db 3473 ATGCAATTTCTGGGTCTATCAGAGTCCAAACTTTCTGACTCTGGCAGCCAGCTCCCTG 3532  
Qy 3421 GATTTTGAAGTTCAGCCTTATGCACTGCTCTCACACTTCTTAATTTTCAAGCTTGA 3480  
Db 3533 GATTTTGAAGTTCAGCCTTATGCACTGCTCTCACACTTCTTAATTTTCAAGCTTGA 3592  
Qy 3481 GGAATCCCAATTATAGAGGCTATAGAGGCAAGAAATAGCTTGGGTGTTTGCATCT 3540  
Db 3593 GGAATCCCAATTATAGAGGCTATAGAGGCAAGAAATAGCTTGGGTGTTTGCATCT 3652  
Qy 3541 ACTCAGATATACCACTGTGGCTTTAAAGCTCTGTCTGATTTTGCAGCCCTATAGTAATCA 3600  
Db 3653 ACTCAGATATACCACTGTGGCTTTAAAGCTCTGTCTGATTTTGCAGCCCTATAGTAATCA 3712  
Qy 3601 GAAAGCAAAATATCCAGATGACCGGTGACGGGCTAGCTCAACAAATCTT----- 3651  
Db 3713 GAAAGCAAAATATCCAGATGACCGGTGACGGGCTAGCTCAACAAATCTTGAAGTTT 3772  
Qy 3652 -----CTTGTGGTACAGCA 3669  
Db 3773 CTGATTTGACACACACACCGCTTACTCTTACAGACAGAGCTTGTGTGATACAGCA 3832  
Qy 3670 ATGCAAGTTAATATTTCCGAAATGTTTGGATTTGCTATTTTGCAGCTCAATGTTGA 3729  
Db 3833 ATGCAAGTTAATATTTCCGAAATGTTTGGATTTGCTATTTTGCAGCTCAATGTTGA 3892  
Qy 3730 TATATTTGAAGGCTTGTGGGTTCTTGAAGACGAGATCTATCCAAATCAAGAACCC 3789  
Db 3893 TATATTTGAAGGCTTGTGGGTTCTTGAAGACGAGATCTATCCAAATCAAGAACCC 3952  
Qy 3750 TTTGATTTAGATTTGCTGTATTAAGAAATTAAGTGAATCTCAATCATGTGATTTGAAT 3849  
Db 3953 TTTGATTTAGATTTGCTGTATTAAGAAATTAAGTGAATCTCAATCATGTGATTTGAAT 4012  
Qy 3850 GTGTATCAAGCTTTCCGGCCCGGGTAGAGTGCATGAGCTTATGGAATTAACCTA 3909  
Db 4013 GTGTATCAAGCTTTCCGGCCCGGGTAGAGTGCATGAGCTTATGGAATTAACCTA 4072  
Qy 3910 TTAAGTGGCTTATGATGCTTTCAGAGCAATTTCTTGAAGAGACAGTGAAGAAAGTG 3969  
Db 4073 TTAAGTGGCTTATGATGCTTTCAGAGCAATTTCTTGAAGAGACAGTGAAGAAAGTG 4132  
Qy 3970 GAATATGATCATGAGAAACTCAACTCTATTTAGATTTCTGTAAATGAAACCAAGTTTGT 4029  
Db 4133 GAATATGATCATGAGAAACTCAACTCTATTTAGATTTCTGTAAATGAAACCAAGTTTGT 4192  
Qy 4030 GTTATATTTCTGCTGTGAGAACTTTAAAGTTTCAATACCAAGATGCTTCAAGTCC 4089  
Db 4133 GTTATATTTCTGCTGTGAGAACTTTAAAGTTTCAATACCAAGATGCTTCAAGTCC 4252  
Qy 4090 ATAAGTGAATTAATATGAGCCAGAGAGACAGGCGGTGAGAAATTAACCTGTAAGTGAAG 4149  
Db 4253 ATAAGTGAATTAATATGAGCCAGAGAGACAGGCGGTGAGAAATTAACCTGTAAGTGAAG 4312

Qy 4150 CTGCTCTCTGAGACCTTTGAGTATGTCAGAGGCTGCGCTCTTGTGAGATGAGCT 4209  
Db 4313 CTGCTCTCTGAGACCTTTGAGTATGTCAGAGGCTGCGCTCTTGTGAGATGAGCT 4312  
Qy 4210 TCAGGCTCCCATCATCTCTTCACTATTTTATTTTCTGTTTCAAGCTTCTGACTTT 4269  
Db 4373 TCAGGCTCCCATCATCTCTTCACTATTTTATTTTCTGTTTCAAGCTTCTGACTTT 4432  
Qy 4270 ATGGAATTTGGCTGTGA 4287  
Db 4433 ATGGAATTTGGCTGTGA 4450

RESULT 6  
AX537610

LOCUS AX537610 4335 bp DNA linear PAT 23-NOV-2002  
DEFINITION Sequence 1 from Patent WO02070738.  
ACCESSION AX537610  
VERSION AX537610.1 GI:25269531  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1  
AUTHORS Schuh, A. and Ouwehand, W.  
TITLE Diagnosis and treatment of blood disorders  
JOURNAL Patent: WO 02070738-A 1 12-SEP-2002;  
Schuh, Andre (CA)

FEATURES  
source  
1. .4335  
Location/Qualifiers

/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
1. .>4335  
/note="unamed protein product"  
/codon\_start=1  
/protein\_id="CAD57496.1"  
/db\_xref="GI:25269532"

/translation="MGSPPLTAAHLACVCTAALAAVAPGRFLVTARGLIRPGAVTI  
GVELLEPCSVTVKAEILKTASNLTVSLAEAGVFEKSGFKTLPLSLPLNSADELY  
ELAVTGRODELIFNSRLSPFKRISVPIQDKALYKQKQKFP1VTLFSFKFY  
KTLNLIKIDKSNLIQOMLSQSDLSITFOLSHPIKIDMSIOVAVDDQYYS.  
FOVEEYLPKFEVTLQTPLYCSKMSKLNITAKYTGKVKQKVDVLTFLPSSFMK  
KQNTKTETKINGSANBSRDESEKNNDSNGLSEYLDISGPEVLEITVTSVGTI  
SRNSTVFRKQHDYITIEFPDYYTLKPSLFTATVKTTRADGQTLBERNNVIT  
VTORNYEWSGNSGNOKEAVKINVTYQSGTFKIEPILDESELQKAYFLS  
KSNMAVSLFKSPKTYIOLKTRDENIKVSPFELVSGNRLKELSYMVNSQOLVA  
VGKNSIMFSLTPENSWTPKACIVIVYLEDGEIISDVLKIPOLVFNKIKLWYSIV  
KAEPSEKSLRISVTPDSIVGIVAVKSVLMAASNDITMENVHELELYNTGYIG  
MFMNSPAVFOCGIWLTDNLTKDYIDGVYDNLAEVAREMEHEGIIYIDHPSLCS  
SPHYKAPPEFTWLDTRMGRTIYQEBVTPDSITSWANPISDILGLITTTTY  
ELQAPQPEFTPLNPIYSVIRGEFALETITFNKADTEVAVIIIEKSDKDIITTSB  
INATGHOQTLIDNPEDATVLPFRPHLGEIPITVATLSPADAVTOMILVABEI  
EKXSOSIILDLINRLQSLTKLTSFSPPTVTSERVOITAGVLSINLAL  
IRMYGCGEONINAPNIVLIDYTKKQTLNKKKLSFMQGOBELYQREDS  
SFSAPGVDPBSGTLSAFVRLCEADPYDIDONVLAHRTYMLKQHSNGSPMP  
GRVHSELQGNKSPVLTATVITSLIGRYQONIDYESIHLSEFSRGSIDNT  
LALITVALSVSGSPKAEALMLTWRAQEGGQFVWSSEKSDSQPSLIDEVA  
YALISHLQFOETGEGIPIMRLSRORSRSGFASDTOTVAKLSFALMTERNA  
IOVTVGSSPSPVKPLIDTNRLILOELAEIAPVMAVNSANGFALCOLVVVN  
IKVAGSSRRRSRSIONOAFDLDVAVKKNKDINVDINVCFSFGPSRSGALMEVNL  
LSGPMVSEASISLSEYTKVYKXDKNLKLYDSYNEOFNCNIPAVNPKYSNODS  
VSIYDIYEPKROAVRSYNSSEVSLSCDLCSDVQCRCEBDASGSHHSVITLFCFK  
LLVPEMELWL"

BASE COUNT 1296 a 837 c 922 g 1280 t  
ORIGIN

Query Match 98.5%; Score 4221.4; DB 6; Length 4335;  
Best Local Similarity 98.8%; Pred. No. 0;  
Matches 4283; Conservative 0; Mismatches 1; Indels 51; Gaps 1;

Qy 1 ATGAGGAGCCACCGCTCTGACCGCGCCGACCTCTCTGCGTGTGACCGCGCGCTG 60  
Db 1 ATGAGGAGCCACCGCTCTGACCGCGCCGACCTCTCTGCGTGTGACCGCGCGCTG 60  
Qy 61 GCGTGGCTCCGCGCTCGGCTTTCGCTGACAGCCGAGGATCAATCAGAGCCGAGGA 120  
Db 61 GCGTGGCTCCGCGCTCGGCTTTCGCTGACAGCCGAGGATCAATCAGAGCCGAGGA 120  
Qy 121 AATGTGACTATTTGGGAGTGTGAGCTTCTGAAACACTGCGCTTCAAGGTGACTGGAAGCG 180  
Db 121 AATGTGACTATTTGGGAGTGTGAGCTTCTGAAACACTGCGCTTCAAGGTGACTGGAAGCG 180  
Qy 181 GAGCTGCTAAGACAGCATCAAACTCTGCTGTCTGTCTCTGAGCAGAAAGACTTTT 240  
Db 181 GAGCTGCTAAGACAGCATCAAACTCTGCTGTCTGTCTCTGAGCAGAAAGACTTTT 240  
Qy 241 GAAAAAGGCTCTTTAAGACACTTACTCTCCATCCTGACTCTGAAACAGTGAAGTGA 300  
Db 241 GAAAAAGGCTCTTTAAGACACTTACTCTCCATCCTGACTCTGAAACAGTGAAGTGA 300  
Qy 301 ATTTATGAGCTAGTGTAAACCGGACGTAACCGAGATGAGATTTTATTTCTAATAGTACC 360  
Db 301 ATTTATGAGCTAGTGTAAACCGGACGTAACCGAGATGAGATTTTATTTCTAATAGTACC 360  
Qy 361 CGCTTATCATTTGAGACCAAGAGATATCTGTCTTCAATCAACAGCAAGGCTTATAC 420  
Db 361 CGCTTATCATTTGAGACCAAGAGATATCTGTCTTCAATCAACAGCAAGGCTTATAC 420  
Qy 421 AAGCCAAAGCAAGAGTGAAGTTTCCGATGTGTATGACTCTTCAATTTTAAGCCTTAC 480  
Db 421 AAGCCAAAGCAAGAGTGAAGTTTCCGATGTGTATGACTCTTCAATTTTAAGCCTTAC 480  
Qy 481 AAAACCTCTTAAACATCTCTAATTAAGAACCCCAATCAAAATTTGAATCCAAAGTGGTTG 540  
Db 481 AAAACCTCTTAAACATCTCTAATTAAGAACCCCAATCAAAATTTGAATCCAAAGTGGTTG 540  
Qy 541 TCACAAACAAAGTATCTTGAGATCATTTCCAAACCTTTTCAAGCTATCTTCCATCAATA 600  
Db 541 TCACAAACAAAGTATCTTGAGATCATTTCCAAACCTTTTCAAGCTATCTTCCATCAATA 600  
Qy 601 CTGTGACCTGTGCTATTCAGATTCAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 660  
Db 601 CTGTGACCTGTGCTATTCAGATTCAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 660  
Qy 661 GTTTCAAGATGATGATTAACCAAAATTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 720  
Db 661 GTTTCAAGATGATGATTAACCAAAATTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 720  
Qy 721 ATGAATTTAAGCATTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780  
Db 721 ATGAATTTAAGCATTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780  
Qy 781 AAGGAGAGCTAACGCTTAATTTTACCTTTTGGGGGAAAGAAAGAAATATTT 840  
Db 781 AAGGAGAGCTAACGCTTAATTTTACCTTTTGGGGGAAAGAAAGAAATATTT 840  
Qy 841 AAAAAACATTTAAGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
Db 841 AAAAAACATTTAAGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
Qy 901 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
Db 901 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
Qy 961 GTTGAATTTTAAACACAGTGAAGATGATGATGATGATGATGATGATGATGATGATGAT 1020  
Db 961 GTTGAATTTTAAACACAGTGAAGATGATGATGATGATGATGATGATGATGATGATGAT 1020  
Qy 1021 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080  
Db 1021 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080  
Qy 1081 AAGCCATCTCTCAACTTCAAGCAGCACTGTGAAGTGAATCTGTGCTGATGCAACCACTG 1140



```
Db 1081 AAGCCATCTCACTTCAACCCACTGGAAGTAATCTGTGATGAGCAACCAACG 1140
Qy 1141 ACTCTTGAAGAAAGAAATATATGATCATTAACAGTACACAGAGAACTATACTGAG 1200
Db 1141 ACTCTTGAAGAAAGAAATATATGATCATTAACAGTACACAGAGAACTATACTGAG 1200
Qy 1201 TACTGAGCGGATCTTAACAGTGAATCAAGAAATGGAAGCTGTTCAAGAAATTAAT 1260
Db 1201 TACTGAGCGGATCTTAACAGTGAATCAAGAAATGGAAGCTGTTCAAGAAATTAAT 1260
Qy 1261 ACTGCCCCCAAGTGGAACTTTTAAGATTGAATTCCTGAGAGATTCCAGTGA 1320
Db 1261 ACTGCCCCCAAGTGGAACTTTTAAGATTGAATTCCTGAGAGATTCCAGTGA 1320
Qy 1321 CTACAGTTGAAGCCCTATTTCTGTGATGAATAAGAGATGAGCATGTCGTT 1380
Db 1321 CTACAGTTGAAGCCCTATTTCTGTGATGAATAAGAGATGAGCATGTCGTT 1380
Qy 1381 AAGTCTCTAGTAAAGCATATCACTCACTCAAAACAGAGATGAATAATTAAGTGG 1440
Db 1381 AAGTCTCTAGTAAAGCATATCACTCAAAACAGAGATGAATAATTAAGTGG 1440
Qy 1441 TCGCTTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
Db 1441 TCGCTTTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
Qy 1501 GATTCAGAGGAGCAGTGTGCTGATGAGAAACAAATTCACAAATGTCCTTTA 1560
Db 1501 GATTCAGAGGAGCAGTGTGCTGATGAGAAACAAATTCACAAATGTCCTTTA 1560
Qy 1561 CCAGAAATTTCTTGAATCTCAAAAGCTGTATTTGTATTAATTAAGATGATG 1620
Db 1561 CCAGAAATTTCTTGAATCTCAAAAGCTGTATTTGTATTAATTAAGATGATG 1620
Qy 1621 GAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
Db 1621 GAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
Qy 1681 CTATATGAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 1740
Db 1681 CTATATGAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 1740
Qy 1741 ACACAGCCTGATCTCATATGATGATGATGATGATGATGATGATGATGATGAT 1800
Db 1741 ACACAGCCTGATCTCATATGATGATGATGATGATGATGATGATGATGATGAT 1800
Qy 1801 GCTCTATATGATATTAACATGGAATGTCATGATGATGATGATGATGATGAT 1860
Db 1801 GCTCTATATGATATTAACATGGAATGTCATGATGATGATGATGATGATGAT 1860
Qy 1861 TATTAATTAAGCATGATGATGATGATGATGATGATGATGATGATGATGAT 1920
Db 1861 TATTAATTAAGCATGATGATGATGATGATGATGATGATGATGATGATGAT 1920
Qy 1921 TTGACAGATGCAAACTCAAGAGATTAATGATGATGATGATGATGATGATGAT 1980
Db 1921 TTGACAGATGCAAACTCAAGAGATTAATGATGATGATGATGATGATGATGAT 1980
Qy 1981 GCTGAGAGTTTATGAGAGAAATGAGACATATGATGATGATGATGATGATGAT 2040
Db 1981 GCTGAGAGTTTATGAGAGAAATGAGACATATGATGATGATGATGATGATGAT 2040
Qy 2041 GGTAGACATCCATATGTCGAAAGATTTTCAAGACATGATGATGATGATGATGAT 2100
Db 2041 GGTAGACATCCATATGTCGAAAGATTTTCAAGACATGATGATGATGATGATGAT 2100
Qy 2101 ATGGGTTACAGATTTTACAGAAATTTGAAGTATCTGATCTGATCTGATCTGAT 2160
Db 2101 ATGGGTTACAGATTTTACAGAAATTTGAAGTATCTGATCTGATCTGATCTGAT 2160
Qy 2161 GTGGCTACTGTTTGTGATCTCTGAGACCTGGGCTTGAACAATCTCAGTGG 2220
Db 2161 GTGGCTACTGTTTGTGATCTCTGAGACCTGGGCTTGAACAATCTCAGTGG 2220

Db 2161 GTGGCTACTGTTTGTGATCTCTGAGACCTGGGCTTGAACAATCTCAGTGG 2220
Qy 2221 GAGCTCAAGCCTTCCAACTTTTCAATTTTGAATCTTCCCTACTGATGATGATGAT 2280
Db 2221 GAGCTCAAGCCTTCCAACTTTTCAATTTTGAATCTTCCCTACTGATGATGATGAT 2280
Qy 2281 GGTGAAGATTTGCTTGGAAATTAATTAATTAATTAATTAATTAATTAATTAAT 2340
Db 2281 GGTGAAGATTTGCTTGGAAATTAATTAATTAATTAATTAATTAATTAATTAAT 2340
Qy 2341 AAGGTAATCTTGAAGAAAGTACAAATTTGATATCTTAATGATCTTCAAGTAAAT 2400
Db 2341 AAGGTAATCTTGAAGAAAGTACAAATTTGATATCTTAATGATCTTCAAGTAAAT 2400
Qy 2401 GCCACAGGCAACACAGACCCCTTGTGCTCAGTGAAGATGGGCAACTGTCCTTT 2460
Db 2401 GCCACAGGCAACACAGACCCCTTGTGCTCAGTGAAGATGGGCAACTGTCCTTT 2460
Qy 2461 CCCATCAGGCAACACATCTGGAGAAATTCCTATCAGATCAAGCTCTTCAACCA 2520
Db 2461 CCCATCAGGCAACACATCTGGAGAAATTCCTATCAGATCAAGCTCTTCAACCA 2520
Qy 2521 GCTTCTGATGCTGTCAACCAAGATTTTGAATTAAGCTGAAGAAATTAATTAAT 2580
Db 2521 GCTTCTGATGCTGTCAACCAAGATTTTGAATTAAGCTGAAGAAATTAATTAAT 2580
Qy 2581 TCACATCATCTTATTAATGATGATGATGATGATGATGATGATGATGATGAT 2640
Db 2581 TCACATCATCTTATTAATGATGATGATGATGATGATGATGATGATGATGAT 2640
Qy 2641 AGTTTCTATTTCTCTCAATTAACAGTGAATGGCAGTGAAGATTCAGATCACTG 2700
Db 2641 AGTTTCTATTTCTCTCAATTAACAGTGAATGGCAGTGAAGATTCAGATCACTG 2700
Qy 2701 GAGATGTTCTTGTGCTTCCATCAATGATGATGATGATGATGATGATGATGAT 2760
Db 2701 GAGATGTTCTTGTGCTTCCATCAATGATGATGATGATGATGATGATGATGAT 2760
Qy 2761 TGTGTGAACAGACATGATTAATTTGCTCAAAATTTTCAATTTGATGATGATGAT 2820
Db 2761 TGTGTGAACAGACATGATTAATTTGCTCAAAATTTTCAATTTGATGATGATGAT 2820
Qy 2821 AAAAAGAAACAACTGACATTAATTTGAAGAAAGCTCTTCAATTAAGAGCAAGT 2880
Db 2821 AAAAAGAAACAACTGACATTAATTTGAAGAAAGCTCTTCAATTAAGAGCAAGT 2880
Qy 2881 TACAGAGAGAACTTCTATCAAGAGAGATGCTCTTCAAGTCTTTGGAAATTAAT 2940
Db 2881 TACAGAGAGAACTTCTATCAAGAGAGATGCTCTTCAAGTCTTTGGAAATTAAT 2940
Qy 2941 GACCTTCTGGAGACATTTGCTGATGATGATGATGATGATGATGATGATGATGAT 3000
Db 2941 GACCTTCTGGAGACATTTGCTGATGATGATGATGATGATGATGATGATGATGAT 3000
Qy 3001 CTTATAGATATTAATGATCAGAAATGTTTACACAGAAATTAATTAATTAATTAAT 3060
Db 3001 CTTATAGATATTAATGATCAGAAATGTTTACACAGAAATTAATTAATTAATTAAT 3060
Qy 3061 CAGAAATCAACGATGAAATTTTGGATCAGAGAGATGATTAATGATGATGATGAT 3120
Db 3061 CAGAAATCAACGATGAAATTTTGGATCAGAGAGATGATTAATGATGATGATGAT 3120
Qy 3121 GGCATTAAGATCCAGTAACTTAACAGCTTAATTTGATGATGATGATGATGATGAT 3180
Db 3121 GGCATTAAGATCCAGTAACTTAACAGCTTAATTTGATGATGATGATGATGATGAT 3180
Qy 3181 AAGTATCAAGCTTAATGATGATGATGATGATGATGATGATGATGATGATGAT 3240
Db 3181 AAGTATCAAGCTTAATGATGATGATGATGATGATGATGATGATGATGATGAT 3240
Qy 3241 AAGGTAATTTACAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3300
Db 3241 AAGGTAATTTACAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3300
```

QY	3301	AGTCTAAAGGAAAGAAAGCTTTGAAATATGCGACTTGGAGAAGCAACAAGAGTGGC	3366
Dp	3301	AGCTTAAGGAAAGAAAGCTTTGAAATATGCGACTTGGAGAAGCAACAAGAGTGGC	3366
QY	3361	ATGCATTTCTGGGTGCATCAGAGTCCAACTTTCGACTCCCTGGCAGCCAGCTCCCTG	3420
Dp	3361	ATGCATTTCTGGGTGCATCAGAGTCCAACTTTCGACTCCCTGGCAGCCAGCTCCCTG	3420
QY	3421	GATATTGAAGTTGCAGCCTATGCACTGCTCTCAGACTTCTTAACAATTTCAAGCTTGTAG	3480
Dp	3421	GATATTGAAGTTGCAGCCTATGCACTGCTCTCAGACTTCTTAACAATTTCAAGCTTGTAG	3480
QY	3481	GGAAATCCCAATTATGAGGTGGCTAGACGAGCAAGAATAATAGCTGGGGTGGTTGCACT	3540
Dp	3481	GGAAATCCCAATTATGAGGTGGCTAGACGAGCAAGAATAATAGCTGGGGTGGTTGCACT	3540
QY	3541	ACTCAGATATCCACTGTGGCTTTTAAAGGCTCTGTGTAATTTGCGCCCTATGAAATACA	3600
Dp	3541	ACTCAGATATCCACTGTGGCTTTTAAAGGCTCTGTGTAATTTGCGCCCTATGAAATACA	3600
QY	3601	GAAAAGCAAAATATCCAAAGTGAACCGTGAACGGGCGCTAGCTCAACAAGTCTT-----	3651
Dp	3601	GAAAAGCAAAATATCCAAAGTGAACCGTGAACGGGCGCTAGCTCAACAAGTCTT-----	3651
QY	3652	-----CTTGCTGTGTGTACAGCA-3653	3653
Dp	3652	-----CTTGCTGTGTGTACAGCA-3653	3653
QY	3661	CTGATTGACACACACAAACCGCTTACTCCTTGACAGACAGACAGCTTGCTGTGTACAGCA	3720
Dp	3661	CTGATTGACACACACAAACCGCTTACTCCTTGACAGACAGACAGCTTGCTGTGTACAGCA	3720
QY	3670	ATGCGAGTTAATATTTCCGCAAAATGGTTTGGATTGGCTATTTTGCACTCAATGTTGTA	3729
Dp	3721	ATGCGAGTTAATATTTCCGCAAAATGGTTTGGATTGGCTATTTTGCACTCAATGTTGTA	3780
QY	3730	TATAATGTGAAGCTTCTGGGTCTTCTGAAGACGAAATCTATCCAAATCAAGAACCC	3789
Dp	3781	TATAATGTGAAGCTTCTGGGTCTTCTGAAGACGAAATCTATCCAAATCAAGAACCC	3840
QY	3790	TTTGATTTAGATGTGTCGTGTAAAGAAATAAAGATGATCTCAATCATGTGGATTGAT	3849
Dp	3841	TTTGATTTAGATGTGTCGTGTAAAGAAATAAAGATGATCTCAATCATGTGGATTGAT	3900
QY	3850	GTGCTGACAAAGCTTTGGGGCCCGGGTGAAGTGCATGGGCTTATGGAAGTTAACTTA	3909
Dp	3901	GTGCTGACAAAGCTTTGGGGCCCGGGTGAAGTGCATGGGCTTATGGAAGTTAACTTA	3960
QY	3910	TTAAGTGCCTTATGTCCTTCAAGAACAAATTTCTGAGCGAGACAGTGAAGAAATG	3968
Dp	3961	TTAAGTGCCTTATGTCCTTCAAGAACAAATTTCTGAGCGAGACAGTGAAGAAATG	4020
QY	3970	GAATATGATCATGAAAACTCAACCTCTATTATTAAGTTCTGTAAATGAACCACATTTGT	4029
Dp	4021	GAATATGATCATGAAAACTCAACCTCTATTATTAAGTTCTGTAAATGAACCACATTTGT	4080
QY	4030	GTTAATATTTCTGCTGTGTGAAGAACTTTAAAGTTCAATATCCCAAGATGCTTCACTGCC	4088
Dp	4081	GTTAATATTTCTGCTGTGTGAAGAACTTTAAAGTTCAATATCCCAAGATGCTTCACTGCC	4140
QY	4090	ATAGTGAATTAATGAGCAAGAGGAGACAGGGGGTGAAGTTAACTCTGAAGTGAAG	4149
Dp	4141	ATAGTGAATTAATGAGCAAGAGGAGACAGGGGGTGAAGTTAACTCTGAAGTGAAG	4200
QY	4150	CTGTCTCTCTGTGACCTTGTGCAGTGATGTCCAGGGCTCGCTCTTGTGAGAGTGAAGT	4209
Dp	4201	CTGTCTCTCTGTGACCTTGTGCAGTGATGTCCAGGGCTCGCTCTTGTGAGAGTGAAGT	4260
QY	4210	TCAGGCTCCCATCATCACTCTTCAAGTCAATTTTATTTCTGTTTCAAGCTTGTGACTTT	4268
Dp	4261	TCAGGCTCCCATCATCACTCTTCAAGTCAATTTTATTTTCTGTTTCAAGCTTGTGACTTT	4320
QY	4270	ATGGAACCTTGGCTG 4284	
Dp	4321	ATGGAACCTTGGCTG 4335	

[illegible]

Db 181 GAGCTGCTAAGACAGATCAAACTCACGTCTCTGCTCGAGACAGAGAGCTTT 240  
Qy 241 GAAAAAGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGAATG 300  
Db 241 GAAAAAGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGAATG 300  
Qy 301 ATTTATGACCTAAGTGTAAACCGACGTACCCAGATGAGATTTTATCTTAATAGTACC 360  
Db 301 ATTTATGACCTAAGTGTAAACCGACGTACCCAGATGAGATTTTATCTTAATAGTACC 360  
Qy 361 CGCTTATCATTTGAGACCAAGAAATATCTGTCTTATTCAAACAGACAGAGCTTATAC 420  
Db 361 CGCTTATCATTTGAGACCAAGAAATATCTGTCTTATTCAAACAGACAGAGCTTATAC 420  
Qy 421 AAGCCAAAGCAAGAGTGAAGTTGSCATTTGTTACACTCTCTCAGATTTTAAGCCTTAC 480  
Db 421 AAGCCAAAGCAAGAGTGAAGTTGSCATTTGTTACACTCTCTCTCAGATTTTAAGCCTTAC 480  
Qy 481 AAAACCTCTTTAAACATTTCTCATTTAAGACCCCAATCAAAATTTGATCCAAAGTGGT 540  
Db 481 AAAACCTCTTTAAACATTTCTCATTTAAGACCCCAATCAAAATTTGATCCAAAGTGGT 540  
Qy 541 TCACAAACAAGTGTATTTGAGTCAATTCGAAAATTTTCCAGTATCTTCCATCCAAAT 600  
Db 541 TCACAAACAAGTGTATTTGAGTCAATTCGAAAATTTTCCAGTATCTTCCATCCAAAT 600  
Qy 601 CTGAGTACGTGCTATTCAGTTCAGGATGACAGACATATATCAATCATTTGAG 660  
Db 601 CTGAGTACGTGCTATTCAGTTCAGGATGACAGACATATATCAATCATTTGAG 660  
Qy 661 GTTTCAGATATGATTTTACCAAAATTTGAAGTACTTTCAGACACATATATTTGTTCT 720  
Db 661 GTTTCAGATATGATTTTACCAAAATTTGAAGTACTTTCAGACACATATATTTGTTCT 720  
Qy 721 ATGATTTCTAAGCTTTTAATAGTACATCAGGCAAGATATCATATGGAAGCCAGT 780  
Db 721 ATGATTTCTAAGCTTTTAATAGTACATCAGGCAAGATATCATATGGAAGCCAGT 780  
Qy 781 AAAAGGAGCTAAGCTTACATTTTATCCTTTATCGTTTGGGAAAGAAAGAAATAT 840  
Db 781 AAAAGGAGCTAAGCTTACATTTTATCCTTTATCGTTTGGGAAAGAAAGAAATAT 840  
Qy 841 ACAAACAATTTAAGATTAATGATCTGCAAACTTCTCTTTAATGATGAAGATGA 900  
Db 841 ACAAACAATTTAAGATTAATGATCTGCAAACTTCTCTTTAATGATGAAGATGA 900  
Qy 901 AATGTAATGATTTCTTCAAAATGATCTTCTGAATACCTGATCTATCTTCCCTGAGCA 960  
Db 901 AATGTAATGATTTCTTCAAAATGATCTTCTGAATACCTGATCTATCTTCCCTGAGCA 960  
Qy 961 GTAGAAATTTTAAACAAGTGAAGATTAATGATTAATTTCAAGAAATGTAAGCAT 1020  
Db 961 GTAGAAATTTTAAACAAGTGAAGATTAATGATTAATTTCAAGAAATGTAAGCAT 1020  
Qy 1021 AATGATTTCTTCAAGCAATGATTTACATCATTTGATTTTGAATTAATCTAATCTGCTG 1080  
Db 1021 AATGATTTCTTCAAGCAATGATTTACATCATTTGATTTTGAATTAATCTAATCTGCTG 1080  
Qy 1081 AAGCATCTCTCAACTTCAAGCACTGAGCACTGAGGATTAATCTGATGATGCAACCACTG 1140  
Db 1081 AAGCATCTCTCAACTTCAAGCACTGAGCACTGAGGATTAATCTGATGATGCAACCACTG 1140  
Qy 1141 ACTCTTGAAGAAAGAAATTAATGATCATTAAGTGAACAGAGAACTTATCTGAG 1200  
Db 1141 ACTCTTGAAGAAAGAAATTAATGATCATTAAGTGAACAGAGAACTTATCTGAG 1200  
Qy 1201 TACTGAGGAGATCTTAACAGTGAATTCAGAAATGGAAGCTTTAGAAAAATTAAT 1260  
Db 1201 TACTGAGGAGATCTTAACAGTGAATTCAGAAATGGAAGCTTTAGAAAAATTAAT 1260  
Qy 1261 ACTGTCCCCCAAGTGAACCTTTTAAGATGAAATTCGAATCCTGAGAGATTCAGTGA 1320  
Db 1261 ACTGTCCCCCAAGTGAACCTTTTAAGATGAAATTCGAATCCTGAGAGATTCAGTGA 1320

Db 1261 ACTGTCCCCCAAGTGAACCTTTTAAGATGAAATTCGAATCCTGAGAGATTCAGTGA 1320  
Qy 1321 CTACAGTTGAAGGCTATTTCTTGTGTATTAAGATGAGATGAGATGATGATCTGTT 1380  
Db 1321 CTACAGTTGAAGGCTATTTCTTGTGTATTAAGATGAGATGAGATGATGATCTGTT 1380  
Qy 1381 AAGTCTCTTGAAGACATACATCACTAATAAACAAGATGAATAATTAAGTGGGA 1440  
Db 1381 AAGTCTCTTGAAGACATACATCACTAATAAACAAGATGAATAATTAAGTGGGA 1440  
Qy 1441 TGGCTTTTGAATGTGTGTATGAGCAACAAGATTAAGAGATTAAGTATGAT 1500  
Db 1441 TGGCTTTTGAATGTGTGTATGAGCAACAAGATTAAGAGATTAAGTATGAT 1500  
Qy 1501 GTATCAGGAGACAGTGTGTGTATGAGCAACAAGATTAAGAGATTAAGTATGAT 1560  
Db 1501 GTATCAGGAGACAGTGTGTGTATGAGCAACAAGATTAAGAGATTAAGTATGAT 1560  
Qy 1561 CCAGAAATTTCTTGAATCTCAGAAAGCTGTGTATTTGATTAATTAAGATGATG 1620  
Db 1561 CCAGAAATTTCTTGAATCTCAGAAAGCTGTGTATTTGATTAATTAAGATGATG 1620  
Qy 1621 GAAATTTAATGATGATTTCTAATAATTTCTGTTCAGCTTTGTTTAAATAATGAT 1680  
Db 1621 GAAATTTAATGATGATTTCTAATAATTTCTGTTCAGCTTTGTTTAAATAATGAT 1680  
Qy 1681 CTATATTTGAAGTGAAGTGAAGCTGAACATCTGAAGAGCTCTTATGATCTCTG 1740  
Db 1681 CTATATTTGAAGTGAAGTGAAGCTGAACATCTGAAGAGCTCTTATGATCTCTG 1740  
Qy 1741 ACACAGCTGATCTCAGATTTGAGATTTGAGCTGTTCAGAAAGTGTATCTGAT 1800  
Db 1741 ACACAGCTGATCTCAGATTTGAGATTTGAGCTGTTCAGAAAGTGTATCTGAT 1800  
Qy 1801 GCTCTAATGATTTTAAACAAGTGAAGATGCTGATGATGATGATGATGATGAT 1860  
Db 1801 GCTCTAATGATTTTAAACAAGTGAAGATGCTGATGATGATGATGATGATGAT 1860  
Qy 1861 TATTAATTTAAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920  
Db 1861 TATTAATTTAAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920  
Qy 1921 TTGACAGATGCAAACTCAGAGATTAATTTGATGATGATGATGATGATGAT 1980  
Db 1921 TTGACAGATGCAAACTCAGAGATTAATTTGATGATGATGATGATGATGAT 1980  
Qy 1981 GCTGAGGATTTATGAGAGAAATGAAGACATATTTGATGATGATGATGAT 2040  
Db 1981 GCTGAGGATTTATGAGAGAAATGAAGACATATTTGATGATGATGATGAT 2040  
Qy 2041 GGTAGCAGTCCACATGCTCCGAAAGATTTTCCAGAGACTGATTTGGCTACACAC 2100  
Db 2041 GGTAGCAGTCCACATGCTCCGAAAGATTTTCCAGAGACTGATTTGGCTACACAC 2100  
Qy 2101 ATGGGTTACAGATTTTCAAGAAATTTGAAGTGAACGTGATCTGATCTGAT 2160  
Db 2101 ATGGGTTACAGATTTTCAAGAAATTTGAAGTGAACGTGATCTGATCTGAT 2160  
Qy 2161 GTGGCTACTGTTTGTGATCTTGAAGACCTGAGTCTTGAACATCACTCAAGT 2220  
Db 2161 GTGGCTACTGTTTGTGATCTTGAAGACCTGAGTCTTGAACATCACTCAAGT 2220  
Qy 2221 GAGCTCCAAAGCTTCCAAACATTTTCAATTTTGAATCTTCCCTAATCTGAT 2280  
Db 2221 GAGCTCCAAAGCTTCCAAACATTTTCAATTTTGAATCTTCCCTAATCTGAT 2280  
Qy 2281 GGTGAAGATTTGCTTGAAGAAATTAATCAATTTTGAAGATGAGTCAAGT 2340  
Db 2281 GGTGAAGATTTGCTTGAAGAAATTAATCAATTTTGAAGATGAGTCAAGT 2340  
Qy 2341 AAGGTAATCATTTGAAGAAATGAACAATTTGATGATGATGATGATGATGAT 2400  
Db 2341 AAGGTAATCATTTGAAGAAATGAACAATTTGATGATGATGATGATGATGAT 2400

QY 2401 GCCACAGCCACACAGACAGACCCCTTCTGGTTCACAGTGAAGATGGGGCAACTGTTCTTTT 2460  
 DB 2401 GCCACAGCCACACAGACAGACCCCTTCTGGTTCACAGTGAAGATGGGGCAACTGTTCTTTT 2460  
 QY 2461 CCCATCAGGCCACACATCTGGAGAGAAATTCCTTACACATCAAGCTCTTCCCACT 2520  
 DB 2461 CCCATCAGGCCACACATCTGGAGAGAAATTCCTTACACATCAAGCTCTTCCCACT 2520  
 QY 2521 GCTTCTGATGCTGCACAGATGATTTAGTAAAGGCTGAAGAAATGAAGAAATCAAT 2580  
 DB 2521 GCTTCTGATGCTGCACAGATGATTTAGTAAAGGCTGAAGAAATGAAGAAATCAAT 2580  
 QY 2581 TCACATCCATCTTATAGACTTGAAGTCAATAGGCTACAGAGTACCCTGAACCTTGG 2640  
 DB 2581 TCACATCCATCTTATAGACTTGAAGTCAATAGGCTACAGAGTACCCTGAACCTTGG 2640  
 QY 2641 AGTTCTCATTTCTCTTAATACAGTGAAGTGAAGAAAGTTCAATCACTGCAAT 2700  
 DB 2641 AGTTCTCATTTCTCTTAATACAGTGAAGTGAAGAAAGTTCAATCACTGCAAT 2700  
 QY 2701 GGAGATGCTTGTGCTTCATCAATGGCTTACGCTTATGATGCTGGAATGCTTATGGC 2760  
 DB 2701 GGAGATGCTTGTGCTTCATCAATGGCTTACGCTTATGATGCTGGAATGCTTATGGC 2760  
 QY 2761 TGTGCTGAACAGAAATGATTAATTTTCTCCAAATATTTACATTTTGGATTTCTGACT 2820  
 DB 2761 TGTGCTGAACAGAAATGATTAATTTTCTCCAAATATTTACATTTTGGATTTCTGACT 2820  
 QY 2821 AAAAAAGAAACACAGATGATTAATTTGAAGAAAAAGCTCTTCAATTTATGAGCAAGT 2880  
 DB 2821 AAAAAAGAAACACAGATGATTAATTTGAAGAAAAAGCTCTTCAATTTATGAGCAAGT 2880  
 QY 2881 TACAGAGAGAACTTCTATACAGAGGAGATGCTTTCAGGCTTTTGGGAATAT 2940  
 DB 2881 TACAGAGAGAACTTCTATACAGAGGAGATGCTTTCAGGCTTTTGGGAATAT 2940  
 QY 2941 GACCTTCTGGAGACATTTGCTGACAGTCTTGTGTTTAAAGATGTTTCTTGAAGCCAT 3000  
 DB 2941 GACCTTCTGGAGACATTTGCTGACAGTCTTGTGTTTAAAGATGTTTCTTGAAGCCAT 3000  
 QY 3001 CCTTACATGATTAATGATCAAGATGCTTGAACAGAAATCACTTGGCTTAAAGACAT 3060  
 DB 3001 CCTTACATGATTAATGATCAAGATGCTTGAACAGAAATCACTTGGCTTAAAGACAT 3060  
 QY 3061 CAGAAATCCACGAGATTTTGGATCAGAGAAAGATTCATAGTCAAGCTTCAAGGT 3120  
 DB 3061 CAGAAATCCACGAGATTTTGGATCAGAGAAAGATTCATAGTCAAGCTTCAAGGT 3120  
 QY 3121 GGCAATAAAGTCCAGTACACTTACAGCTATATTTGAATCTTCTCTGGGATATAGA 3180  
 DB 3121 GGCAATAAAGTCCAGTACACTTACAGCTATATTTGAATCTTCTCTGGGATATAGA 3180  
 QY 3181 AAGTATCAGCTTAATGATGTCAGAGATCTATCAATTTTGGAGTCTGAATTCAGT 3240  
 DB 3181 AAGTATCAGCTTAATGATGTCAGAGATCTATCAATTTTGGAGTCTGAATTCAGT 3240  
 QY 3241 AGAGGAATTTCAAGCAATTAATCTTACGCTTATTAATCTTATGCAATGTCAGTGGG 3300  
 DB 3241 AGAGGAATTTCAAGCAATTAATCTTACGCTTATTAATCTTATGCAATGTCAGTGGG 3300  
 QY 3301 AGTCTTAAAGGAGAAAGCTTTGAAATATGCTGAAGTGAAGGAGAAAGAAAGTGGC 3360  
 DB 3301 AGTCTTAAAGGAGAAAGCTTTGAAATATGCTGAAGTGAAGGAGAAAGAAAGTGGC 3360  
 QY 3361 ATGCAATTTCTGGGTCTATCAGAGTCCAACTTTCTGACTCTGGACCCAGCTCCCTG 3420  
 DB 3361 ATGCAATTTCTGGGTCTATCAGAGTCCAACTTTCTGACTCTGGACCCAGCTCCCTG 3420  
 QY 3421 GATATTTGAAGTTGACAGCTTATGCACTGCTTCAACATTTCAAGATTTCTGAG 3480  
 DB 3421 GATATTTGAAGTTGACAGCTTATGCACTGCTTCAACATTTCAAGATTTCTGAG 3480

QY 3481 GAAATCCCAATTAATGAGTGGCTTAAGCAGGCAAGAAATAGCTTGGGTGTTTGCATCT 3540  
 DB 3481 GAAATCCCAATTAATGAGTGGCTTAAGCAGGCAAGAAATAGCTTGGGTGTTTGCATCT 3540  
 QY 3541 ACTCAGATTAATGAGTGGCTTAAAGGCTCTGTCTGAATTTTGAAGCCCTTAATGAATCA 3600  
 DB 3541 ACTCAGATTAATGAGTGGCTTAAAGGCTCTGTCTGAATTTTGAAGCCCTTAATGAATCA 3600  
 QY 3601 GAAAGCAAAATATCAAGTGAAGCCGTGACGGGGCTGATCCAAATGCTT----- 3651  
 DB 3601 GAAAGCAAAATATCAAGTGAAGCCGTGACGGGGCTGATCCAAATGCTT----- 3651  
 QY 3652 -----CTTGTGTGTGTAACGCCA 3669  
 DB 3652 -----CTTGTGTGTGTAACGCCA 3669  
 QY 3661 CTGATGACACACACACCGCTTACTCTTACAGACAGAGCTTGTGTGTGTAACGCCA 3720  
 DB 3661 CTGATGACACACACACCGCTTACTCTTACAGACAGAGCTTGTGTGTGTAACGCCA 3720  
 QY 3720 ATGCAATTAATTTTCCGCAAAATGTTTGGATTTGCTATTTGCTCAAGTGAATGTA 3780  
 DB 3720 ATGCAATTAATTTTCCGCAAAATGTTTGGATTTGCTATTTGCTCAAGTGAATGTA 3780  
 QY 3781 TATATGTAAGGCTTCTGGGTCTTCTAGAAGACAGATCTATCCAAATCAAGAAAGCC 3840  
 DB 3781 TATATGTAAGGCTTCTGGGTCTTCTAGAAGACAGATCTATCCAAATCAAGAAAGCC 3840  
 QY 3790 TTTGATTTAGATGTTGCTGTAAAGAAATAAAGATGATCTCAATCATGTGATTTGAAT 3849  
 DB 3790 TTTGATTTAGATGTTGCTGTAAAGAAATAAAGATGATCTCAATCATGTGATTTGAAT 3849  
 QY 3841 TTTGATTTAGATGTTGCTGTAAAGAAATAAAGATGATCTCAATCATGTGATTTGAAT 3900  
 DB 3841 TTTGATTTAGATGTTGCTGTAAAGAAATAAAGATGATCTCAATCATGTGATTTGAAT 3900  
 QY 3901 GTGTGTCAAGCTTTTGGGCGCGGTGAGAGTGGCATGTGCTTTATGAAATTAACCTTA 3960  
 DB 3901 GTGTGTCAAGCTTTTGGGCGCGGTGAGAGTGGCATGTGCTTTATGAAATTAACCTTA 3960  
 QY 3910 TTAAGTGGCTTTATGAGTGGCTTCAAGAGCAATTTCTGAGACGAGACAGTGAAGAAAGTG 3969  
 DB 3910 TTAAGTGGCTTTATGAGTGGCTTCAAGAGCAATTTCTGAGACGAGACAGTGAAGAAAGTG 3969  
 QY 3961 TTAAGTGGCTTTATGAGTGGCTTCAAGAGCAATTTCTGAGACGAGACAGTGAAGAAAGTG 4020  
 DB 3961 TTAAGTGGCTTTATGAGTGGCTTCAAGAGCAATTTCTGAGACGAGACAGTGAAGAAAGTG 4020  
 QY 3970 GAAATGATCATGGAAGAACTCAACCTTATTAATGATCTGTAATGAAGAAAGCAAGTTTGT 4029  
 DB 3970 GAAATGATCATGGAAGAACTCAACCTTATTAATGATCTGTAATGAAGAAAGCAAGTTTGT 4029  
 QY 4021 GATATGATCATGGAAGAACTCAACCTTATTAATGATCTGTAATGAAGAAAGCAAGTTTGT 4080  
 DB 4021 GATATGATCATGGAAGAACTCAACCTTATTAATGATCTGTAATGAAGAAAGCAAGTTTGT 4080  
 QY 4030 GTTAAATTTCTGTGTGAGAAATTTTAAAGTTTCAATACCAAGATGCTTCAAGTGTCC 4089  
 DB 4030 GTTAAATTTCTGTGTGAGAAATTTTAAAGTTTCAATACCAAGATGCTTCAAGTGTCC 4089  
 QY 4081 GTTAAATTTCTGTGTGAGAAATTTTAAAGTTTCAATACCAAGATGCTTCAAGTGTCC 4140  
 DB 4081 GTTAAATTTCTGTGTGAGAAATTTTAAAGTTTCAATACCAAGATGCTTCAAGTGTCC 4140  
 QY 4090 ATAGTGAATTAATGAGCCAGAGAGCAGCGGTGAGAGTTCAACTCTGAAGTGAAG 4149  
 DB 4090 ATAGTGAATTAATGAGCCAGAGAGCAGCGGTGAGAGTTCAACTCTGAAGTGAAG 4149  
 QY 4141 ATAGTGAATTAATGAGCCAGAGAGCAGCGGTGAGAGTTCAACTCTGAAGTGAAG 4200  
 DB 4141 ATAGTGAATTAATGAGCCAGAGAGCAGCGGTGAGAGTTCAACTCTGAAGTGAAG 4200  
 QY 4150 CTGTCTCTGTGACCTTTGACAGTATGTCAGGGCTGCCGCTCTGTGAGATGAGCT 4209  
 DB 4150 CTGTCTCTGTGACCTTTGACAGTATGTCAGGGCTGCCGCTCTGTGAGATGAGCT 4209  
 QY 4201 CTGTCTCTGTGACCTTTGACAGTATGTCAGGGCTGCCGCTCTGTGAGATGAGCT 4260  
 DB 4201 CTGTCTCTGTGACCTTTGACAGTATGTCAGGGCTGCCGCTCTGTGAGATGAGCT 4260  
 QY 4210 TCAGGCTCCATCATCACTCTTCAAGTCAATTTTATTTTCTGTTCAAGCTTCTGACTTT 4269  
 DB 4210 TCAGGCTCCATCATCACTCTTCAAGTCAATTTTATTTTCTGTTCAAGCTTCTGACTTT 4269  
 QY 4261 TCAGGCTCCATCATCACTCTTCAAGTCAATTTTATTTTCTGTTCAAGCTTCTGACTTT 4320  
 DB 4261 TCAGGCTCCATCATCACTCTTCAAGTCAATTTTATTTTCTGTTCAAGCTTCTGACTTT 4320  
 QY 4270 ATGGAATTTGGGCTG 4284  
 DB 4270 ATGGAATTTGGGCTG 4284  
 QY 4321 ATGGAATTTGGGCTG 4335  
 DB 4321 ATGGAATTTGGGCTG 4335

RESULT 8  
 AX534942  
 LOCUS AX534942 3535 bp DNA linear PAT 22-NOV-2002  
 DEFINITION Sequence 9 from Patent W002070696.  
 ACCESSION AX534942  
 VERSION AX534942.1 GI:25261490  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1

AUTHORS Schuh, A. and Sutherland, R.D.  
TITLE Cdi09 nucleic acid molecules, polypeptides and methods of use  
JOURNAL Patent: WO 02070696-A 9 12-SEP-2002;

Schuh, Andre (CA) ; Sutherland, Robert D. (CA)

FEATURES  
source Location/Qualifiers

1..3535

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

CDS

113..3316

/note="unnamed protein product"

/codon\_start=1

/protein\_id="CAD57235.1"

/db\_xref="GI:25261491"

/translation="MGGPLLTAAHLICVCTAALAVAPRPFLVTAAGIIRPGNVTI  
GVELEHCPQVTVAKELTKTASNLTVSLAEAGPEKESFETLTPSLASADIEY  
ELRTGRDDELFSNSTRLSPETKISVFIOTDPAKYKPEUKVRYTLTPKDEPY  
KTSNLILKIDPKSNLIQWLSQSDLGIVSKTFOLSHNLIIDWMSIQVANDQTYOS  
FOVSEYVLPKFEVTLQTLPCSMNSHNLGITTAKYKPKVKGDLTFLPLTWGK  
KNKITKPKINGSANFSFNDEBKVMVDSNGLSEYLDLSPGPELITTVESVTGI  
SHNVSTNVEFKQHDYIIEFFDYTVLKPENFTATKVRADGNOLTEERNNVIT  
VTORNTYETWSSGNSGNOMKAEVOKINTVPOSGPKIEPFLIDSSSELQKAYFLGS  
KSMAYHSLFKSPKTYIOLKTRDENIKVSGPELVYSGNKKLKEISVWVRQOLVA  
VKKNSIMSELTPENSWTPKACIVITYIIEDDEIISDVLTIPVQVFNKIKLWYKLV  
KAEPEKSVLRISVQPSI VGI VAVDKSVNLNMSNDITMENVAHELWNTGYLGS  
MENSFPAVEQECGLWLTNDANLTKDYIIDGVYNAEYAEERFMEENHGLVIDHFDLGS  
SPHNRHPEPEITWLDNMGYRIYQEFETVDPDSITSWATGVFISEDLGLITTPV  
ELAORPPEIPLNLPYSVIRGEPEALEITFNVLKDAPEKVKIIEKSDKPELITLME  
INATGHOCTLIPSEDDGATVLPPIRPHLGEIPITYTALSPASDADTOMILVKEGI  
EKSYISQSLDLIDNRKLOSTLTKLSFSPENPVTGSEVQITADIVLSPSINGLASL  
IRMPYCGGQNMNINPAPNTYIIDLVLTKKQQLNDNLKEXLSPMGQYQYSELVYOREDG  
SFSAFGNYDPSGSNTWLSAFVLRCLFEADYIDIDONVLRITYTWMKGHOKSNGEEMDP  
GRVISELQGGKSPVTLTAYIVTSLGVRKYQVFRIR"

BASE COUNT 1108 a 675 c 731 g 1021 t  
ORIGIN

Query Match 74.4%; Score 3187.4; DB 6; Length 3535;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3188; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 ATGCGAGGCGCCACCGTCTCTGACCGCGCCACCTCTCTGCGTGTGACCGCGCGCTG 60  
113 ATGCGAGGCGCCACCGTCTCTGACCGCGCCACCTCTCTGCGTGTGACCGCGCGCTG 172  
61 GCCGTGGCTCCCGGGGCTCGGTTCTGCGACAGCCCGAGGATATATGAGCCCGAGGA 120  
173 GCCGTGGCTCCCGGGGCTCGGTTCTGCGACAGCCCGAGGATATATGAGCCCGAGGA 232  
121 AATGTGACTATTTGGGTGAGGCTTCTGAAACACTGCGCTTCAAGGTGACTGTGAAGCG 180  
223 AATGTGACTATTTGGGTGAGGCTTCTGAAACACTGCGCTTCAAGGTGACTGTGAAGCG 292  
181 GAGCTGCTCAAGACAGCATCAAACTCACTGCTCTGCTGCGAAGCAGAAAGAGCTTTT 240  
293 GAGCTGCTCAAGACAGCATCAAACTCACTGCTCTGCTGCGAAGCAGAAAGAGCTTTT 352  
241 GAAAAAGGCTTTTAAAGACACTTACTCTTCATCACTACTCTGAAAGCAGTGAAGTG 300  
353 GAAAAAGGCTTTTAAAGACACTTACTCTTCATCACTACTCTGAAAGCAGTGAAGTG 412  
301 ATTATGAGCTACGTGTAAACCGGACGTACCCAGAGTGAATTTATTTCTTAATAGTACC 360  
413 ATTATGAGCTACGTGTAAACCGGACGTACCCAGAGTGAATTTATTTCTTAATAGTACC 472  
361 CGCTTATCATTTTGAAGCAAGAGATATCTGTCTTCAATCAACAGACAGAGCCTTATAC 420  
473 CGCTTATCATTTTGAAGCAAGAGATATCTGTCTTCAATCAACAGACAGAGCCTTATAC 532  
421 AAGCCAAAGCAAGAGTGAAGTTTGGCATTTGTTTACCTTCTCAGATTTTTAAAGCTTAC 480  
533 AAGCCAAAGCAAGAGTGAAGTTTGGCATTTGTTTACCTTCTCAGATTTTTAAAGCTTAC 592  
481 AAAACCTTTTAAACATCTCATTAAGAGCCCAATCAAAATTTGATCCAAAGTGTG 540

593 AAAACCTTTTAAACATCTCATTAAGAGCCCAATCAAAATTTGATCCAAAGTGTG 652  
541 TCACAAACAAGGATCTTGGAGTCAATTCACAAAATCTTTCAGCTCTCCCATCCATA 600  
653 TCACAAACAAGGATCTTGGAGTCAATTCACAAAATCTTTCAGCTCTCCCATCCATA 712  
601 CTGTGTGATCGTCTTATCAAGTTCAGTGAATGACACAGACATATATCATATTTAG 660  
713 CTGTGTGATCGTCTTATCAAGTTCAGTGAATGACACAGACATATATCATATTTAG 772  
661 GTTTCAGATATATATATCAAAATTTGAGTGAATGACCTTTGACAGACCATTAATTTG 720  
773 GTTTCAGATATATATATCAAAATTTGAGTGAATGACCTTTGACAGACCATTAATTTG 832  
721 AAGAAATTCATACATTTAAATGATGACATACAGGAAAGTATACATATGGAAGCAG 780  
833 ATGAATTCATACATTTAAATGATGACATACAGGAAAGTATACATATGGAAGCAG 892  
781 AAGAGAGAGTAAAGCTTACATTTTACCTTATCTTTGGGAAAGAAATAATTT 840  
893 AAGAGAGAGTAAAGCTTACATTTTACCTTATCTTTGGGAAAGAAATAATTT 952  
841 ACAAACAATTTAAGATTAATGATCTGCAAACTTCTCTTTTAATGATGAAGATGAA 900  
953 ACAAACAATTTAAGATTAATGATCTGCAAACTTCTCTTTTAATGATGAAGATGAA 1012  
901 AATGTAAATGATCTTCAAAATGATCTTCTGAAATCTGATCTATCTTCCCTGAGCA 960  
1013 AATGTAAATGATCTTCAAAATGATCTTCTGAAATCTGATCTATCTTCCCTGAGCA 1072  
961 GTAGAAATTTTAAACACAGTGAACAGATCAGTAAAGGATTTCAAGAAATGAAGCACT 1020  
1073 GTAGAAATTTTAAACACAGTGAACAGATCAGTAAAGGATTTCAAGAAATGAAGCACT 1132  
1021 AATGTGTTCTTCAACCAATGATTTACATCATTTGAGTTTGTATTAATCTATCTGCTG 1080  
1133 AATGTGTTCTTCAACCAATGATTTACATCATTTGAGTTTGTATTAATCTATCTGCTG 1192  
1081 AAGCATCTCTCAATTTCAACAGCCCTGTAAGAGTAACTCTGCTGATGGAACCAACTG 1140  
1193 AAGCATCTCTCAATTTCAACAGCCCTGTAAGAGTAACTCTGCTGATGGAACCAACTG 1252  
1141 ACTCTGAAGAAAGAAATTAATGATGATCAATCAAGTGAACAGAGAACTATCTGAG 1200  
1253 ACTCTGAAGAAAGAAATTAATGATGATCAATCAAGTGAACAGAGAACTATCTGAG 1312  
1201 TACTGAGGAGATCTTAAGTGAATCAAGAAATGAGAGCTGTTCAAGAAATTAATTTAT 1260  
1313 TACTGAGGAGATCTTAAGTGAATCAAGAAATGAGAGCTGTTCAAGAAATTAATTTAT 1372  
1261 ACTGTCCCCCAAGTGAACCTTTTAAGATGTAATCCAAATCTGAGAGATTCAGTGA 1320  
1373 ACTGTCCCCCAAGTGAACCTTTTAAGATGTAATCCAAATCTGAGAGATTCAGTGA 1432  
1321 CTACAGTGAAGGCTATTTCTTGTGTGTAAGTAAAGTACATGAGGCTCATAGTGTGTT 1380  
1433 CTACAGTGAAGGCTATTTCTTGTGTGTAAGTAAAGTACATGAGGCTCATAGTGTGTT 1492  
1381 AAGTCTCTAGTAAGACATACATCAACTTAAAAAACAAGATGAATAATTAAGTGGGA 1440  
1493 AAGTCTCTAGTAAGACATACATCAACTTAAAAAACAAGATGAATAATTAAGTGGGA 1552  
1441 TCGCTTTTGAAGTGTGTGTAAGTGAACAAAGATTTGAAGATTAATGATGTA 1500  
1553 TCGCTTTTGAAGTGTGTGTAAGTGAACAAAGATTTGAAGATTAATGATGTA 1612  
1501 GTATCAGGAGGACATGTTGTGCTGTAGAGAAACAAATTTCAACAAATGTTCTCTTAACA 1560  
1613 GTATCAGGAGGACATGTTGTGCTGTAGAGAAACAAATTTCAACAAATGTTCTCTTAACA 1672  
1561 CCAGAAATTTCTGAGTCCAAAGCTGTGTAATTTGTAATTAATGAAGATGATGG 1620

Db 1673 CCAGAAATCTTGAGCTCCAAAAGCTGTGTAAATGTGTATTAATTAAGATGATGCG 1732  
 Qy 1621 GAATTAATTAAGTATGTTCTAAAAAATTCCTGTGTAGCTTTGTTTAAAAATTAAGTAAAG 1680  
 Db 1733 GAATTAATTAAGTATGTTCTAAAAAATTCCTGTGTAGCTTTGTTTAAAAATTAAGTAAAG 1792  
 Qy 1681 CTATTAATTAAGTATGTTCTAAAAAATTCCTGTGTAGCTTTGTTTAAAAATTAAGTAAAG 1740  
 Db 1793 CTATTAATTAAGTATGTTCTAAAAAATTCCTGTGTAGCTTTGTTTAAAAATTAAGTAAAG 1852  
 Qy 1741 ACAGAGCTGTAGCTATGTTGAGTATGTAGCTGTGCAAAAGTGAATCTGATGAAT 1800  
 Db 1853 ACAGAGCTGTAGCTATGTTGAGTATGTAGCTGTGCAAAAGTGAATCTGATGAAT 1912  
 Qy 1801 GCTCTATGATTAATTAAGTATGTTGAGTATGTAGCTGTGCAAAAGTGAATCTGATGAAT 1860  
 Db 1913 GCTCTATGATTAATTAAGTATGTTGAGTATGTAGCTGTGCAAAAGTGAATCTGATGAAT 1972  
 Qy 1861 TATTAATTAAGTATGTTGAGTATGTAGCTGTGCAAAAGTGAATCTGATGAAT 1920  
 Db 1973 TATTAATTAAGTATGTTGAGTATGTAGCTGTGCAAAAGTGAATCTGATGAAT 2032  
 Qy 1921 TTGACAGTGAATTAAGTATGTTGAGTATGTAGCTGTGCAAAAGTGAATCTGATGAAT 1980  
 Db 2033 TTGACAGTGAATTAAGTATGTTGAGTATGTAGCTGTGCAAAAGTGAATCTGATGAAT 2092  
 Qy 1981 GCTGAGAGTATTAAGTATGTTGAGTATGTAGCTGTGCAAAAGTGAATCTGATGAAT 2040  
 Db 2093 GCTGAGAGTATTAAGTATGTTGAGTATGTAGCTGTGCAAAAGTGAATCTGATGAAT 2152  
 Qy 2041 GGTAGAGTATTAAGTATGTTGAGTATGTAGCTGTGCAAAAGTGAATCTGATGAAT 2100  
 Db 2153 GGTAGAGTATTAAGTATGTTGAGTATGTAGCTGTGCAAAAGTGAATCTGATGAAT 2212  
 Qy 2101 ATGAGTATTAAGTATGTTGAGTATGTAGCTGTGCAAAAGTGAATCTGATGAAT 2160  
 Db 2213 ATGAGTATTAAGTATGTTGAGTATGTAGCTGTGCAAAAGTGAATCTGATGAAT 2272  
 Qy 2161 GTGAGTATTAAGTATGTTGAGTATGTAGCTGTGCAAAAGTGAATCTGATGAAT 2220  
 Db 2273 GTGAGTATTAAGTATGTTGAGTATGTAGCTGTGCAAAAGTGAATCTGATGAAT 2332  
 Qy 2221 GAGTATTAAGTATGTTGAGTATGTAGCTGTGCAAAAGTGAATCTGATGAAT 2280  
 Db 2333 GAGTATTAAGTATGTTGAGTATGTAGCTGTGCAAAAGTGAATCTGATGAAT 2392  
 Qy 2281 GGTAGAGTATTAAGTATGTTGAGTATGTAGCTGTGCAAAAGTGAATCTGATGAAT 2340  
 Db 2393 GGTAGAGTATTAAGTATGTTGAGTATGTAGCTGTGCAAAAGTGAATCTGATGAAT 2452  
 Qy 2341 AAGTATTAAGTATGTTGAGTATGTAGCTGTGCAAAAGTGAATCTGATGAAT 2400  
 Db 2453 AAGTATTAAGTATGTTGAGTATGTAGCTGTGCAAAAGTGAATCTGATGAAT 2512  
 Qy 2401 GCCAGAGTATTAAGTATGTTGAGTATGTAGCTGTGCAAAAGTGAATCTGATGAAT 2460  
 Db 2513 GCCAGAGTATTAAGTATGTTGAGTATGTAGCTGTGCAAAAGTGAATCTGATGAAT 2572  
 Qy 2461 CCCAGAGTATTAAGTATGTTGAGTATGTAGCTGTGCAAAAGTGAATCTGATGAAT 2520  
 Db 2573 CCCAGAGTATTAAGTATGTTGAGTATGTAGCTGTGCAAAAGTGAATCTGATGAAT 2632  
 Qy 2521 GCTTCTGATGTTGAGTATGTTGAGTATGTAGCTGTGCAAAAGTGAATCTGATGAAT 2580  
 Db 2633 GCTTCTGATGTTGAGTATGTTGAGTATGTAGCTGTGCAAAAGTGAATCTGATGAAT 2692  
 Qy 2581 TCAGATTAAGTATGTTGAGTATGTAGCTGTGCAAAAGTGAATCTGATGAAT 2640  
 Db 2693 TCAGATTAAGTATGTTGAGTATGTAGCTGTGCAAAAGTGAATCTGATGAAT 2752  
 Qy 2641 AGTTTCTGATGTTGAGTATGTTGAGTATGTAGCTGTGCAAAAGTGAATCTGATGAAT 2700  
 Db 2753 AGTTTCTGATGTTGAGTATGTTGAGTATGTAGCTGTGCAAAAGTGAATCTGATGAAT 2812

Qy 2701 GGAGATGTTCTTGCTCTTCATCAATGCTTAGCTCATGATGATGCTTATGCT 2760  
 Db 2813 GGAGATGTTCTTGCTCTTCATCAATGCTTAGCTCATGATGATGCTTATGCT 2872  
 Qy 2761 TGTGTGTAACAAGATGATTAATTTTGTCTCAAAATTTTACATTTTGTATGCTACT 2820  
 Db 2873 TGTGTGTAACAAGATGATTAATTTTGTCTCAAAATTTTACATTTTGTATGCTACT 2932  
 Qy 2821 AAAAAGAAACAATGATTAATTTTGAAGAAAAGCTCTTCATTTATGAGCAAGT 2880  
 Db 2933 AAAAAGAAACAATGATTAATTTTGAAGAAAAGCTCTTCATTTATGAGCAAGT 2992  
 Qy 2881 TACCAGAGAACTTCTTATCAAGAGGAAAGATGCTCTTTCAGTGTGTTGGAATAT 2940  
 Db 2993 TACCAGAGAACTTCTTATCAAGAGGAAAGATGCTCTTTCAGTGTGTTGGAATAT 3052  
 Qy 2941 GACCTTCTGAGAGCACTTGCTGTGAGCTTTGTTTAAAGATTTTCCCTTGAAGCCGAT 3000  
 Db 3053 GACCTTCTGAGAGCACTTGCTGTGAGCTTTGTTTAAAGATTTTCCCTTGAAGCCGAT 3112  
 Qy 3001 CCTTACATGATTAATGATCAGAAATGTTTACAGAAATATACATGCTTGAAGCAAT 3060  
 Db 3113 CCTTACATGATTAATGATCAGAAATGTTTACAGAAATATACATGCTTGAAGCAAT 3172  
 Qy 3061 CAGAAATCCAAAGCTGATTAATTTGGGATCCAGAAAGATGATTAAGTGAAGCTTCAAGT 3120  
 Db 3173 CAGAAATCCAAAGCTGATTAATTTGGGATCCAGAAAGATGATTAAGTGAAGCTTCAAGT 3232  
 Qy 3121 GGCATTAAGATCCAGTAATCACTTACAGCTTATTTGTAATCTTCTCCCTGGATATAGA 3180  
 Db 3233 GGCATTAAGATCCAGTAATCACTTACAGCTTATTTGTAATCTTCTCCCTGGATATAGA 3292  
 Qy 3181 AAGTATCAG 3189  
 Db 3293 AAGTATCAG 3301

RESULT 9  
 AX534944  
 LOCUS 3535 bp DNA linear PAT 22-NOV-2002  
 DEFINITION Sequence 11 from Patent WO02070696.  
 ACCESSION AX534944  
 VERSION AX534944.1 GI:25261494  
 KEYWORDS  
 SOURCE  
 ORGANISM Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1  
 Schuh, A. and Sutherland, R.D.  
 AUTHORS  
 TITLE Gd109 nucleic acid molecules polypeptides and methods of use  
 JOURNAL Patent: WO 02070696-A 11 12-SEP-2002;  
 Schuh, Andre (CA) ; Sutherland, Robert D. (CA)  
 FEATURES  
 Location/Qualifiers  
 1..3535  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 113..3316  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAD57236.1"  
 /db\_xref="GI:25261495"  
 /translation="MGGPPLTAAHLICVCTAALAAVAPGRPLVTAIGIIRPGAVTI  
 GVELLEHCBQVTVKALELKTASNLTVSLEAGVEKESFKTLPLSLPNSADELY  
 ELAVRTQDEILFNSSTRISPEKRIISVFIQDKALYRKQKQKRIIVLFDPKFY  
 KTSILNLIQDPEKSNLIQWLSQSDGLSVISKTQDSLSHPILGWSLOVQNDQTYOS  
 FOVEYVLLPFKEVTLQTPLYCSMNSKHLNGTIRAKTYGKPVAGDVTLPPLPSFMGK  
 KNITTEPKIKNSANSFNDSEKHMVDSNGLSSEYLDLSSPGVEILLTVTSYVGI  
 SRVSTNVNPFKQHDYIIEFPDYTVLKPSLNPATYKATRAQDNOULIERBNVYIT  
 VQRNATETWSGNSGNQKMEAVQKINVTVPQSGTRKIEFPILDESSSELDKAYPLGS  
 KSMVAHSLFKSPSKTYIQLKTRDENIKVSSPEELVAVSNKRLKELSYMVVSRGOLVA



VGKQNSTPESLTPENSWPKACVIVYIEDGEIISDVLIKIPVQVFNKIKLYWSKV  
 KAESEKSLRISTVOPDSIVGIIVAVDKSVNLNANSDIEMENVHEIYNTGYLGS  
 MEMNSFAVEOEGLWVLDNANIKDYIGBVNANASAEITMENVEHGLIDIDFELGS  
 SHVAKHPETWIMLDJNMGSRIOBEFETVPDSITSWATPFVISEDGLGTLTPV  
 ELIAROPPEIPLNLPYSVRGEPALEITIENYLDARBNVKIIEKSDKEDLMTSNE  
 INATHOQTLVPSDEGATVLEPIRPTHGELPIVTTALSPISPAADVOMIIVKAEI  
 EKSYQSILIDJDNRLQSTLTSFSPPNVTGSEKQIATAIDVGLPISNGIASL  
 IRMPYCGEOMNINFAPNYIIDLVLTKKQLTNDNKEKALFMRGQYRELLYQREDS  
 SEFAFNGDPSGSTWLSAFVLCFLEADYIDIDONVLFRTYTMKKGHKSNGEPMWD  
 GRVISELGGKSPVTLTAYIVTSLIGRKQYVFR1"

BASE COUNT 1107 a 676 c 731 g 1021 t  
 ORIGIN

Query Match 74.3%; Score 3185.8; DB 6; Length 3535;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 3187; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 ATGCAGGGCCACCGCTCTGACCGCGCCACCTCTCGGTGTGACCGCGCGCTG 60  
 113 ATGCAGGGCCACCGCTCTGACCGCGCCACCTCTCGGTGTGACCGCGCGCTG 172  
 61 GCGTGGCTCCCGGCGCTGCTGTGTGTGACAGCCCGAGATCATGAGCCGAGGA 120  
 173 GCGTGGCTCCCGGCGCTGCTGTGTGTGACAGCCCGAGATCATGAGCCGAGGA 232  
 121 AATGTACTATTGGGGTGGAGCTTCTGAAACACTGCGCTTCAAGGTGACTGAAAGCG 180  
 233 AATGTACTATTGGGGTGGAGCTTCTGAAACACTGCGCTTCAAGGTGACTGAAAGCG 292  
 181 GAGCTGCTCAAGACAGCATCAACTGCTCTGTCTGTCTGTGAGAGAGAGAGCTTT 240  
 293 GAGCTGCTCAAGACAGCATCAACTGCTCTGTCTGTCTGTGAGAGAGAGAGCTTT 352  
 241 GAAAAAGGCTTTTAAAGACCTTACCTTCAATCACTCTCTGAAAGAGTGAATGAG 300  
 353 GAAAAAGGCTTTTAAAGACCTTACCTTCAATCACTCTCTGAAAGAGTGAATGAG 412  
 301 AATTATGAGCTACGTGTAAACCGGACGTACCCAGAGTGAATTTATTTCTAATAGTACC 360  
 413 AATTATGAGCTACGTGTAAACCGGACGTACCCAGAGTGAATTTATTTCTAATAGTACC 472  
 361 GCGTTATCTTTGAGACCAAGAGAAATATCTGTCTTCAATCAACAGACAGGCTTATAC 420  
 473 GCGTTATCTTTGAGACCAAGAGAAATATCTGTCTTCAATCAACAGACAGGCTTATAC 532  
 421 AAGCCAAAGCAAGAGTGAAGTTGGCATTTGTTACACTTTCTCAGATTTTAAAGCTTAC 480  
 533 AAGCCAAAGCAAGAGTGAAGTTGGCATTTGTTACACTTTCTCAGATTTTAAAGCTTAC 592  
 481 AAAAACTCTTAAACCTCTCAATTAAGACCCCAATCAATTTGATCCAAAGGTG 540  
 593 AAAAACTCTTAAACCTCTCAATTAAGACCCCAATCAATTTGATCCAAAGGTG 652  
 541 TCACAAACAAGTATCTTGGAGTCAATTTCCAAAATTTTCACTATCTTCCATCCATA 600  
 653 TCACAAACAAGTATCTTGGAGTCAATTTCCAAAATTTTCACTATCTTCCATCCATA 712  
 601 CTTCGTGACTGCTTATTTAAGTTCAAGTGAATGACACAGACATTTATCATCATTTTAC 660  
 713 CTTCGTGACTGCTTATTTAAGTTCAAGTGAATGACACAGACATTTATCATCATTTTAC 772  
 661 GTTTCAGAAATATGTATTAACAAATTTGAAAGTGAAGTTCGACAGCCATTAATGTTCT 720  
 773 GTTTCAGAAATATGTATTAACAAATTTGAAAGTGAAGTTCGACAGCCATTAATGTTCT 832  
 721 ATGAATTTCAACATTTTAAATGTAACATCAGCGCAAGATATACATATGGAAGCCAGTG 780  
 833 ATGAATTTCAACATTTTAAATGTAACATCAGCGCAAGATATACATATGGAAGCCAGTG 892  
 781 AAGAGAGAGTAAAGCTTACATTTTAACTTTATCTTTTGGGAAAGAAAGAAATAT 840  
 893 AAGAGAGAGTAAAGCTTACATTTTAACTTTATCTTTTGGGAAAGAAAGAAATAT 952

841 AAAAAACATTTAAGATTAATGATCTGCAAACTTCTTTTAAATGATGAAGATGAAA 900  
 953 AAAAAACATTTAAGATTAATGATCTGCAAACTTCTTTTAAATGATGAAGATGAAA 1012  
 901 AATGTAATGATTTCTTCAAAATGAGACTTTTGTGAATCTGTGATCTATCTTTCCCTGAGACA 960  
 1013 AATGTAATGATTTCTTCAAAATGAGACTTTTGTGAATCTGTGATCTATCTTTCCCTGAGACA 1072  
 961 GTAGAAATTTTAAACAGAGTGAAGCAATCAGTTACAGTATTTTGAAGAAATGATGACT 1020  
 1073 GTAGAAATTTTAAACAGAGTGAAGCAATCAGTTACAGTATTTTGAAGAAATGATGACT 1132  
 1021 AATGTGTTCTTCAAGCAACATGATTAACATGAGTTTGTGATTTTGAATTAATACTGCTTGG 1080  
 1133 AATGTGTTCTTCAAGCAACATGATTAACATGAGTTTGTGATTTTGAATTAATACTGCTTGG 1192  
 1081 AAGCATCTCTCAACTTCAAGCAACATGATTAACATGAGTTTGTGATTTTGAATTAATACTGCT 1140  
 1193 AAGCATCTCTCAACTTCAAGCAACATGATTAACATGAGTTTGTGATTTTGAATTAATACTGCT 1252  
 1141 ACTCTGAAAGAAAGAAATATGATGATCAATTAACATGACACAGAGAACTATTAATGAG 1200  
 1253 ACTCTGAAAGAAAGAAATATGATGATCAATTAACATGACACAGAGAACTATTAATGAG 1312  
 1201 TACTGAGCGGATCTTAACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1260  
 1313 TACTGAGCGGATCTTAACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1372  
 1261 ACTGTCCCCCAAGTGAAGCTTTTAAAGTGAATTTCCCAATCTGTGAGATTCAGTGAAG 1320  
 1373 ACTGTCCCCCAAGTGAAGCTTTTAAAGTGAATTTCCCAATCTGTGAGATTCAGTGAAG 1432  
 1321 CTACAGTTGAAGGCTTATTTCTGTGTGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1380  
 1433 CTACAGTTGAAGGCTTATTTCTGTGTGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1492  
 1381 AAGTCTCTAGTGAAGACATACATCACTTAATAAACAAGATGAAGTGAAGTGAAGTGAAG 1440  
 1493 AAGTCTCTAGTGAAGACATACATCACTTAATAAACAAGATGAAGTGAAGTGAAGTGAAG 1552  
 1441 TCGCTTTTGAAGTGGTGTAGTGAAGCAAAAGATTTGAAGATTAATGATGA 1500  
 1553 TCGCTTTTGAAGTGGTGTAGTGAAGCAAAAGATTTGAAGATTAATGATGA 1612  
 1501 GTATCCAGGGACAGTGGTGGCTGTGAGAAAACAAAATTTCAACATGTTCTCTTTAACA 1560  
 1613 GTATCCAGGGACAGTGGTGGCTGTGAGAAAACAAAATTTCAACATGTTCTCTTTAACA 1672  
 1561 CGAGAAATTTCTGGAGCTCAAAAAGCTGTGATTTGATTAATTTGAAGATGATGG 1620  
 1673 CGAGAAATTTCTGGAGCTCAAAAAGCTGTGATTTGATTAATTTGAAGATGATGG 1732  
 1621 GAAATTTAATGATGATTTCTAATAAATTTCTGTGATGATTTTAAATTAATGATTAAG 1680  
 1733 GAAATTTAATGATGATTTCTAATAAATTTCTGTGATGATTTTAAATTAATGATTAAG 1792  
 1681 CTATATTTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1740  
 1793 CTATATTTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1852  
 1741 ACACAGCTGACTCATGTTGGAGTTGTAGCTGTGACAAAGTGAAGTGAAGTGAAGTGAAGTGA 1800  
 1853 ACACAGCTGACTCATGTTGGAGTTGTAGCTGTGACAAAGTGAAGTGAAGTGAAGTGAAGTGA 1912  
 1801 GCTCTTAATGATTAATTAACATGAGAAATGTGTGATCAAGTGAAGTGAAGTGAAGTGAAGTGA 1860  
 1913 GCTCTTAATGATTAATTAACATGAGAAATGTGTGATCAAGTGAAGTGAAGTGAAGTGAAGTGA 1972  
 1861 TATTTATTTGAGCAGTGTATGAATTTCTTTGACAGTCTTTTCAAGAAATGTGAGTCTGGGTA 1920  
 1973 TATTTATTTGAGCAGTGTATGAATTTCTTTTCAAGTCTTTTCAAGAAATGTGAGTCTGGGTA 2032  
 1921 TTGACAGATGCAAACTCAGCAAGATTAATTTGATGATGATGATGATGATGATGATGATGAT 1980

D	2033	TTGACAGATGCAAACTTCACGAGAGATTATGATGCTTTATGACAACTGAGAAAT	2092	D	3113	CTTACATAGATATGATCAGAAATGTTTACACAGAAACATACCTGGCTTAAAGACAT	3172
Q	1981	GCTGAGAGGTTTATGAGAGAAATGAAAGACATATGATGATATTCATGACTTTCTTTG	2040	Q	3061	CAGAAATCCAACGGTGAATTTTGGATCCAGAGAGATTCATAGAGCTTCAAGGT	3120
D	2093	GCTGAGAGGTTTATGAGAGAAATGAAAGACATATGATGATATTCATGACTTTCTTTG	2152	D	3173	CAGAAATCCAACGGTGAATTTTGGATCCAGAGAGATTCATAGAGCTTCAAGGT	3232
Q	2041	GATGAGAGTCCACATGTCGGAAGACATTTTCAGAGACTTGATGATTTGGCTAGACCAAC	2100	Q	3121	GGCAATMAAGTCCAGTACACTTACAGCTTATATTTGTAACCTTCTCTGGATATAGA	3180
D	2153	GATGAGAGTCCACATGTCGGAAGACATTTTCAGAGACTTGATGATTTGGCTAGACCAAC	2212	D	3233	GGCAATMAAGTCCAGTACACTTACAGCTTATATTTGTAACCTTCTCTGGATATAGA	3292
Q	2101	ATGGGTTACAGAGATTTTACCAAGATTTGAAGTAACTGACTGATTTCTATCACTTTGG	2160	Q	3181	AAGTATCAG 3189	
D	2213	ATGGGTTACAGAGATTTTACCAAGATTTGAAGTAACTGACTGATTTCTATCACTTTGG	2272	D	3293	AAGTATCAG 3301	
Q	2161	GTGGCTACTGTTTGTGATCTCTGAGAGACCTGGGCTTTGGACTTAAACAACACTCCAGT	2220	RESULT 10			
D	2273	GTGGCTACTGTTTGTGATCTCTGAGAGACCTGGGCTTTGGACTTAAACAACACTCCAGT	2332	AY083458	4735 bp	mRNA	linear ROD 08-Apr-2002
Q	2221	GAGCTCCAAAGCCTTCCAAACATTTTTCATTTTGTGATCTTCCCTACTCTGTTATGAG	2280	LOCUS	AY083458		
D	2333	GAGCTCCAAAGCCTTCCAAACATTTTTCATTTTGTGATCTTCCCTACTCTGTTATGAG	2392	DEFINITION	Mus musculus GPI-anchored alpha-2 macroglobulin-related protein		
Q	2281	GGTGAAGATTTGCTTTGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	2340	ACCESSION	AY083458	mRNA, complete cds.	
D	2393	GGTGAAGATTTGCTTTGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	2452	VERSION	AY083458.1	GI:20070079	
Q	2341	AAGGTAATCATTTGAAAGAGCAAAATTTGATTTTCATGATCTTCAAGTAATAAT	2400	KEYWORDS	Mus musculus (house mouse)		
D	2453	AAGGTAATCATTTGAAAGAGCAAAATTTGATTTTCATGATCTTCAAGTAATAAT	2512	SOURCE	Mus musculus		
Q	2401	GCCAAGGCAACAGACAGACCTTCTGGTCCAGTGAAGTGGGCAACTGTTCTTTT	2460	ORGANISM	Mus musculus (house mouse)		
D	2513	GCCAAGGCAACAGACAGACCTTCTGGTCCAGTGAAGTGGGCAACTGTTCTTTT	2512	REFERENCE	Hashimoto, M., Ichihara, M. and Takahashi, M.		
Q	2461	CCCATCAGGCCAACAACATCTGGAGAAATTCCTATCAGATCAGACGCTCTTTCACCACT	2520	REFERENCE	Hashimoto, M., Ichihara, M. and Takahashi, M.		
D	2573	CCCATCAGGCCAACAACATCTGGAGAAATTCCTATCAGATCAGACGCTCTTTCACCACT	2632	REFERENCE	Hashimoto, M., Ichihara, M. and Takahashi, M.		
Q	2521	GCTTCTGATGCTGTCACCCAGATGATTTTAAAGGCTGAAGAAATGAATAATCATAT	2580	REFERENCE	Hashimoto, M., Ichihara, M. and Takahashi, M.		
D	2633	GCTTCTGATGCTGTCACCCAGATGATTTTAAAGGCTGAAGAAATGAATAATCATAT	2692	REFERENCE	Hashimoto, M., Ichihara, M. and Takahashi, M.		
Q	2581	TCACATTCATCTTATTAAGCTTGAATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA	2640	REFERENCE	Hashimoto, M., Ichihara, M. and Takahashi, M.		
D	2693	TCACATTCATCTTATTAAGCTTGAATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA	2752	REFERENCE	Hashimoto, M., Ichihara, M. and Takahashi, M.		
Q	2641	AGTTTCTCATTTCTCTTAATACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA	2700	REFERENCE	Hashimoto, M., Ichihara, M. and Takahashi, M.		
D	2753	AGTTTCTCATTTCTCTTAATACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA	2812	REFERENCE	Hashimoto, M., Ichihara, M. and Takahashi, M.		
Q	2701	GGAAATGTTCTTGTGCTTCATCAATGCTTACCTCATGATTCGATGCCCTTAAGGC	2760	REFERENCE	Hashimoto, M., Ichihara, M. and Takahashi, M.		
D	2813	GGAAATGTTCTTGTGCTTCATCAATGCTTACCTCATGATTCGATGCCCTTAAGGC	2872	REFERENCE	Hashimoto, M., Ichihara, M. and Takahashi, M.		
Q	2761	TGTGTGAACAGAACATGATTAATTTGCTCCAAATATTTACATTTTGGATTAATCTGACT	2820	REFERENCE	Hashimoto, M., Ichihara, M. and Takahashi, M.		
D	2873	TGTGTGAACAGAACATGATTAATTTGCTCCAAATATTTACATTTTGGATTAATCTGACT	2932	REFERENCE	Hashimoto, M., Ichihara, M. and Takahashi, M.		
Q	2821	AAAAAGAAACAACTGACAGATTAATTTGAAAGAAAGCTCTTTCATTTATGAGGCAAGT	2880	REFERENCE	Hashimoto, M., Ichihara, M. and Takahashi, M.		
D	2933	AAAAAGAAACAACTGACAGATTAATTTGAAAGAAAGCTCTTTCATTTATGAGGCAAGT	2992	REFERENCE	Hashimoto, M., Ichihara, M. and Takahashi, M.		
Q	2881	TACAGAGAGAACTTCTCTATACAGAGAGAGATGCTCTTTCATGCTTTGGGAATTAAT	2940	REFERENCE	Hashimoto, M., Ichihara, M. and Takahashi, M.		
D	2993	TACAGAGAGAACTTCTCTATACAGAGAGAGATGCTCTTTCATGCTTTGGGAATTAAT	3052	REFERENCE	Hashimoto, M., Ichihara, M. and Takahashi, M.		
Q	2941	GACCTTCTGGAGACATTTGTTGCTTGTGTTTAAAGTGTTCCTTGAAGCCGAT	3000	REFERENCE	Hashimoto, M., Ichihara, M. and Takahashi, M.		
D	3053	GACCTTCTGGAGACATTTGTTGCTTGTGTTTAAAGTGTTCCTTGAAGCCGAT	3112	REFERENCE	Hashimoto, M., Ichihara, M. and Takahashi, M.		
Q	3001	CTTACATAGATATGATCAGAAATGTTTACACAGAAACATACCTGGCTTAAAGACAT	3060	REFERENCE	Hashimoto, M., Ichihara, M. and Takahashi, M.		

D	3113	CTTACATAGATATGATCAGAAATGTTTACACAGAAACATACCTGGCTTAAAGACAT	3172			
Q	3061	CAGAAATCCAACGGTGAATTTTGGATCCAGAGAGATTCATAGAGCTTCAAGGT	3120			
D	3173	CAGAAATCCAACGGTGAATTTTGGATCCAGAGAGATTCATAGAGCTTCAAGGT	3232			
Q	3121	GGCAATMAAGTCCAGTACACTTACAGCTTATATTTGTAACCTTCTCTGGATATAGA	3180			
D	3233	GGCAATMAAGTCCAGTACACTTACAGCTTATATTTGTAACCTTCTCTGGATATAGA	3292			
Q	3181	AAGTATCAG 3189				
D	3293	AAGTATCAG 3301				
FEATURES						
source	1. 4735					
	/organism="Mus musculus"					
	/mol_type="mRNA"					
	/db_xref="taxon:10090"					
	/cell_line="NIH3T3"					
CDS	343..4671					
	/note="GARP"					
	/codon_start=1					
	/product="GPI-anchored alpha-2 macroglobulin-related protein"					
	/protein_id="AA000021.1"					
	/db_xref="GI:20070080"					
	/translation="MRSRLKLSAHLCLCAVLAAGSRFLVTPGRIIRPGANVTIGVLLNSPPOVLVKAQVFKIASNKSRSILIEAGVHGHFTLVLPALPLSPSADKIYE LHIINGSSENIETVSRRTLTRESKISVLTIGTDAFKPKQOEYVFRVLTLCSDLPKPR TSVDYFIDPKSNVYQCFPSQKQDLGVSKTFQLSNMPRIDGWSIQVQVNDQOYSF QVLEVLPRFEVTVQPLVYCSKQQLNGSVIATKTKGPKVSLPLPLPSMGAK KNITKSFENIGPANSFDPNEMKQVMVLPKLTVDSESYENVDSFPGPAIILATYLE SLTGISRMASITVFPKQDYLIEIFDYTVLKPINFATVKSRSQDGLTQEBELEN DLVTVTQKNNHPSORSQDOMDYIOTVNTYIPONGIKIEFPVMSISGLQIKAFLP DGTSSVTWHSMTSPSKTYIOLKTRDEYTKGSPFLMWSGNRQFOLSYMVISKGL VAAKQSSRTSRLTPEASMAKRACTIAYIADESEIINDLIKLPVQVLEPNKAKLPSS KPTVPRSDVSRISATQSDSLVGIIVADKSVLTMEANSISITMETWHELELNTETYL LCFPMNSFAPVQECGLWVLTDAFLIRDSIDEVYDTEYSERFPAENMAVDEPDASS VANVHAKNPFETWIMLVAMVSGSKIEYEFVTVDSITSWASAFVISEDLGQLTPR PAELAFQDFPLFNLPSYVIRGEPELLEVSIVNVLQDTIKVLVLEBSDFDLMTS NTNNTRIKTYQVPRNDGTVLVRPIKPTGELPIVTVASPTASDAVQTVIGKEE GLEKYSKSLDLDTDSNVESKQOQSNRSPFPTVTSERVOITADIGLSINGLS SLIRMPYGGCEQNMITYFAPNTYILDYLTQKQLVNLKEKALSTMROGYKRELIYORE DGSFSAFGIDSSGSTWLSAFVRLCFLEADYIIDDDVLRITYTMANAKKKEGEPW EGRVYHSELQGGTKSPVTLFAYIVTSVIGKQOPNDVDSIKLFEPSRSISDNQI YTLATISYALSVSGPKAEALNLTLMORSEKQPTMLSGPALSGMWQPRSYDIEI AAYALILARTLHHVBSGIPYVEMWLIQQRSLQGPSTQDPTVYALKALSEFALYKENT DIQVLTGTGGLPRSIHFRIDQNLFLHQEELHALDPIYVNVSAHSGGFALQQLNDVY					

D 3113 CTTACATAGATATGATCAGAAATGTTTACACAGAAACATACCTGGCTTAAAGACAT 3172  
 Q 3061 CAGAAATCCAACGGTGAATTTTGGATCCAGAGAGTGAATTCATAGTACGTTCAAGGT 3120  
 D 3173 CAGAAATCCAACGGTGAATTTTGGATCCAGAGAGTGAATTCATAGTACGTTCAAGGT 3232  
 Q 3121 GGCAATMAAGTCCAGTACACTTACAGCTATATTTGTAACCTTCTCTGGATATAGA 3180  
 D 3233 GGCAATMAAGTCCAGTACACTTACAGCTATATTTGTAACCTTCTCTGGATATAGA 3292  
 Q 3181 AAGTATCAG 3189  
 D 3293 AAGTATCAG 3301

RESULT 10  
 AY083458 4735 bp mRNA linear ROD 08-Apr-2002  
 LOCUS AY083458  
 DEFINITION Mus musculus GPI-anchored alpha-2 macroglobulin-related protein  
 ACCESSION AY083458  
 VERSION AY083458.1 GI:20070079  
 KEYWORDS  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 4735)  
 Hashimoto, M., Ichihara, M. and Takahashi, M.  
 Cloning and characterization of GPI-anchored alpha-2 macroglobulin-related protein  
 Unpublished  
 2 (bases 1 to 4735)  
 Hashimoto, M., Ichihara, M. and Takahashi, M.  
 Direct Submission  
 Submitted (11-MAR-2002) Department of Pathology, Nagoya University Graduate School of Medicine, 65 Tsurumai-cho, Showa-ku, Nagoya, Aichi 466-8550, Japan  
 Location/Qualifiers  
 1..4735  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /cell\_line="NIH3T3"  
 343..4671  
 /note="GARP"  
 /codon\_start=1  
 /product="GPI-anchored alpha-2 macroglobulin-related protein"  
 /protein\_id="FAM00021.1"  
 /db\_xref="GI:20070080"  
 /translation="MRSRLLSAHLICCAVALAAGSRFLVTARGIIRPNANTIGVLENSPPQVLKQVPRFKIAKSKSISLEABVFRRGRFRTLVPLALPSBMDKYLHINSGENSEIVSNRTRLPFSKISISLIDQDKAYKQKQKRVLLTCSLDKRYR TSVDIRIKDPSNVLIQWPSOKGDLGVSKTQGLSNPIFGWMSIQVQNDQYQSFQVLEAYLPKFEVYQTPLYCSLQKLSVLAQYGRGVKSGLSLTLPLSFMSCK KNTKSFBEVNPANFSDVSEYEMKVNLLPLDVSSEYENDSPSPGAEIATATE SLTGSRMASTNVPRKQNDYIEIPYTLVLRKSNFTAVVRSRDNGLPTEEN DLVTVVTRKNNPESQRDQNDYIOTVATTPONKTIKEFPWSISSELDQKAYFL DGSSTVSMFTSPSKTYIQKTRREYIKVSPFLWVSGNQKQFQDLSIYWSKQL VVAGKQSRFTSLTPASWAPKCIAYIAYABGELINDILKIPVLEENKRLKWS KPLVKSDDKSLRISATQSDSLVGLIYAVDSKTLAMNSISIMETVHELELVNTY LGFMNPSFAVDECGLMVLTDAITLIDDSIDEVYDRESEYERPEBEANMLUPEDSS VNNVHRKNFPEPMITLDAVMSKTYEEREVTPDSITSWASAPVISEDIGGLTTPAELOAFQPFPLNLPYSVIRGEERFALVSNVTKDITIKVYILIBSDSPILMTS NDNGITVIRKTVQVPRDNGVTIVFPIKPHLGEIPVTVAASPTASDAVDTIVTPE SLEKSYKSVLIDLIDSINVESKQSRFPSPPTVIGSERVQITAGDILSGSINGLS GLIRMYGGEQOMIYFAPNITVILDTYKOKOLVNLKKAALSYMGQYQRELLYORE DGSFSAFGDIDSSGTLWSAFVLRGLADYIDIDODVHRTYTWLMAKKEGEMF EBRVTHSELQGTSPVTLTAYIVSVGYKRYQDNIVQDSIKFLPEFSGISDN YTLAITSVALSTVSGSKAEALNLMQRSEKSDTQFWLSSGSPALSGSNQPSVDLEI AAVALAHTLHVHSEGIPIWMRVLIQDRNSLGEVSVQDPTVALKALSEFSAVHKENT DIOITVTPGIPRISIHFRIDSONLFLHQBELHALDPIITVNVSAHSGFALCOLNDY



NTKGSSSKRRSTENOEYFDJDIVYNNEDIJSHLNLNVCISHLSERTGNTLMEVNL  
 LSGFSASDSIPLSETLKKEYDNGLNLYDSVNESQCVNIPVRYDKVNSIDGS  
 VSEMDYSEPRQAVRSYNTQVTKSSCYLSPDNCKSHTDGATDSLRSSSLIVFCSVL  
 LTFVQH"

BASE COUNT 1338 a 1034 c 1087 g 1276 t  
 ORIGIN

Query Match 63.2%; Score 2709.2; DB 10; Length 4735;  
 Best Local Similarity 77.9%; Pred. No. 0;  
 Matches 3393; Conservative 0; Mismatches 873; Indels 90; Gaps 7;

QY 1 ATGCAGGCGCCACCGCTCTGACCGCCGACCTCTTGGGTGACCGCCGCTG 60  
 DB 343 ATGCGGAGTCGAGGCTCTGACGCGGCCACCTTCTGCTGGCGCGGCTG 402  
 QY 61 GCGGTGCTCCCGGCGCTGCTTGTGTGACAGCCCGAGGATCATCGCCCGAGGA 120  
 DB 403 GCC--GCCCGAGGCTCCGCTTGTGTGACAGCCCGAGGATCATCGCCGAGGA 459  
 QY 121 AATGTACTATTGGGGTGGAGCTTCTGAAACATGCTGCTTCAAGGTGACTGTGAAGCG 180  
 DB 460 AATGTACTATTGGGGTGGATCTCTGAAAAATAGCCCCCAGAGTCTTGTAAAGCT 519  
 QY 181 GAGCTGCTCAAGACAGCATCAACCTCACTGCTCTGCTGAGAGCAGAGAGCTTT 240  
 DB 520 CAGGTGTTCAAGATGACTTCAACAATGAGATCCATCTGGAAACAGAGAGCTTT 579  
 QY 241 GAAAAAGGCTTTTAAAGACCTACTCTTCACTACTACTCTGTAAGAGTGAAGTG 300  
 DB 580 CACAGAGGCAATTTCAAGACTCTGCTTCTCGGCACTACTCTGAGAGTGAAGTGA 639  
 QY 301 ATTATGAGCTACGTGTAACCGGACGTACCCAGAGATGATTTATCTTAATAGTACC 360  
 DB 640 ATTATGAGCTACACATTAAGGACGAATGAGAAATGAGATGATTTCTTCAACGAGACA 699  
 QY 361 CGCTTATCATTTGAGACCAAGAGATATCTGCTTCAATCAACAGACAGAGCTTATAC 420  
 DB 700 CGTTTAACTTTGAGAGCAAGATATCTGCTTCAATCAACAGACAGAGCTTATAC 759  
 QY 421 AAGCCAAAGCAAGATGAGATTTGCAATTTGCACTCTTCAATTTTAAAGCTTAC 480  
 DB 760 AAGCCAAAGCAAGATGAGATTTGCGGCTGCACTCTGCTCAAGCTTAT 819  
 QY 481 AAAACCTTTTAAACCTTCAATTTAGAGACCCCAATCAATTTTCAACAGTGTG 540  
 DB 820 AAGACATCCGTGACATTTTCAATTAAGATCCGAATCTAATGTATCCAAAGTGTG 879  
 QY 541 TCACAAACAAGTATCTTGAAGTCAATTTCCAAAATCTTCAAGTATCTTCCATCCATA 600  
 DB 880 TCTCAAAAAGTATCTTGAAGTCAATTTCCAAAATCTTCAAGTATCTTCAATCCATA 939  
 QY 601 CTGTGAGCTGCTTCAATTTCAAGTCAAGTATGACAGACATTTATCAATCATTTGAG 660  
 DB 940 TTTGTGAGCTGCTTCAATTTCAAGTCAAGTATGACAGACATTTATCAAGCATTTGAG 999  
 QY 661 GTTTCAGATATGATTTTACAAAAATTGAAGTCACTTGCAGACACCATTTATTTGTTCT 720  
 DB 1000 GTCTTGAATATGTGTATCCAAAAGTTCGAAGTCAAGTCAAGCCGCTGATTTGTTC 1059  
 QY 721 ATGAATTTGAAGTATTAATGATCAATCAAGCAAGATATCAATATGGAAGCCAGTG 780  
 DB 1060 TTGAAGTCAAAAGTATTAATGATCAATCAAGCAAGATATCAATATGGAAGCCAGTG 1119  
 QY 781 AAGAGGAGATGACGCTTACATTTTATCTTATCTTTTGGGAAAGAAATATTT 840  
 DB 1120 AAGGGAAGTCTGCTTCAATTTTATCTTGTCTTTTGGGAAAGAAATATTT 1179  
 QY 841 ACAAATAATTAAGATATGATCTGCAAACTTCTTTTATGATGAGATGAAT 900  
 DB 1180 ACAAATAATTTGAGATTAATGATTTGCAAACTTCTTTTATGATGAGATGAAT 1239  
 QY 901 AATGTAAT-----GATTTCTTAATGAGCTTTCTGAATATCTGAT 942

DB 1240 AAGTAATGAATTTGAAACCGCTTACGAGTCTTCGGAAGGAGATTATGAGATGTGAC 1299  
 QY 943 CTATCTTCCCTGACAGCATGATAATTTTAAACAGAGTACAGATCATGATACAGATAT 1002  
 DB 1300 CGGTCTTCTTCTGGGCGGAGATTTATAGTACTGTGAGGAGATCTACAGAGTATC 1359  
 QY 1003 TCAGAAAATGAAGCATATGTGTCTTCAAGCAACATGATTTACATATGAGTTTTT 1062  
 DB 1360 TCAAGATGAGCAGACCAATGTGTTTTTCAAGCAACATGATTTACATATGAGTTTTT 1419  
 QY 1063 GATTATCTACTGTGTTGAAGCATCTTCACTTCAAGCACTGTGAAGTAACTGTT 1122  
 DB 1420 GATTATGAGCTGTGTTAAACATCTCTCAACTTCAAGCACTGTGAAGTAACTGTT 1479  
 QY 1123 GCTATGAGCAACCACTGCTTGAAGAAAGAAATTAATGATGATATACAGTGA 1182  
 DB 1480 TGTGATGCAATCTACTCTGAGAAATTAAGAAATGATTTGTTACAGTATGACC 1539  
 QY 1183 CAGAGAACTATCTGAGTACTGAGCGGATCTAACAGTGAATTCAGAAAATGAGAGCT 1242  
 DB 1540 CAGCGAAAATTAACCACTGAGAG-----TCAGAGAGATCAGAAATGAGATTAAT 1590  
 QY 1243 GTTCAAAAATTAATTAATTAATCTGCTCCCAAGTGAACCTTTAAGATTAATCCCAATC 1302  
 DB 1591 ATCCAGACCGTAATTAATTAATTAATCTGCTCCCAAGTGAATTAATTAATTAATTAAT 1650  
 QY 1303 CTGAGAGATTCAGAGTCAAGTGAAGGCTTCTGTTGATGATTAATTAATTAATTAAT 1362  
 DB 1651 ATGTCATTTCCGCGAGCTCACTGAGAGGCTTATTTCTGATGATTAATTAATTAATTAAT 1710  
 QY 1363 GCAGTTCATTAATCTGTTTAACTCTCTGTAAGATCATCATCACTAATAAACAAGAT 1422  
 DB 1711 ACAATTCACAGATATTAATCTCTCTGTAAGAGTCACTCACTTAATAACAGAGAT 1770  
 QY 1423 GAAATTAATAAGTGGATTCGCTTTGAGTGTGTTAGTGTGAGCAACAAGATTAAG 1482  
 DB 1771 GAATTAATAAGTGGGCTCACTTTGATTTGATGATGATGATGATGATGATGATGATGAT 1830  
 QY 1483 GAGTTAAGCTATTAATGATTAATCAAGGAGAGATGATGATGATGATGATGATGATGAT 1542  
 DB 1831 GACTTAAGCTATTAATGATTAATCAAGGAGAGATGATGATGATGATGATGATGATGATGAT 1890  
 QY 1543 ACAATGTTCTCTTAAACACAGAAAATCTTGAATCTCAAAAAGCTGTGATTTGAT 1602  
 DB 1891 AGAATTTCTCTTAAACACAGAGGCTTCTGAGGCTCAAAAAGCTGTGATTTGAT 1950  
 QY 1603 TATATTAAGATGAGGAAAATTAATTAAGTATGATTTCTTAAATTTCTGTTCAAGTGT 1662  
 DB 1951 TATATTAAGATGAGGAAAATTAATTAAGTATGATTTCTTAAATTTCTGTTCAAGTGT 2010  
 QY 1663 TTTTAAATTAAGATTAAGATTAATTAAGTATGATTAAGTATGATTTCTTAAATTTCTGTT 1722  
 DB 2011 TTTTAAATTAAGATTAAGATTAATTAAGTATGATTAAGTATGATTTCTTAAATTTCTGTT 2070  
 QY 1723 TCTCTTGAAGTCTGTTGACACAGGCTGATCTCAATGATTTGAGATTTGATTTGATTTGAT 1782  
 DB 2071 TCTCTGAGATCTGTTGACACAGGCTGATCTCAATGATTTGAGATTTGATTTGATTTGAT 2130  
 QY 1783 AGTGAATCTGATGAATGCTCTTAATGATTAATCAATGAGAAATGATGATGATGATGAT 1842  
 DB 2131 AGTGAATCTGATGAATGCTCTCAACAGTATTAATCAATGAGAAATGATGATGATGATGAT 2190  
 QY 1843 GAACTTTTAAACAGGATATTTATTTAGGATGATTTGATTTGATTTGATTTGATTTGAT 1902  
 DB 2191 GAACTTTTAAACAGGATATTTATTTAGGATGATTTGATTTGATTTGATTTGATTTGAT 2250  
 QY 1903 GAACTGAGCTCTGAGTATTTGACAGATCAACCTCAAGAAATTAATTAATGATTTGATTTGAT 1962  
 DB 2251 GAATGATGCTCTGAGTATTTGACAGATCAACCTCAATTAAGATTAATGATTTGATTTGAT 2310  
 QY 1963 TATGACAAATGACAAATATGCTGAGAGTTTATGAGAGAAATGAGAGCATATTTGAT 2022  
 DB 2311 TACGATACGAAAGTATTTGAGAGCTTTGACAGGAGAAATGAGGCAAACTGATGAT 2370

QY	2023	ATTGATGACTTTTCTTTGGGTAGAGCTCCAGATGTCGGAAGAACTTTTCCAGAGACTTGG	2082
Db	2371	TTTGAAGATGCTTCTTCAGTTAAACATGTACATGTGCAGAAAGAAATTTTCCAGAAACATGG	2430
QY	2083	ATTTGGCTAGACACCAACAATGGGTTACACAGATTATCCAAAGATTTTGAAGTAACGTACCT	2142
Db	2431	ATTTGGCTAGACGCTTACATGGGATCTTAAGATTTTATGAAGAAATTTGAAGTTACTGTACT	2490
QY	2143	GATTCTATCATCTTTGGGTGGCTACTGGTTTGTGATCTCGAGGACCTGSGTCTTGGGA	2202
Db	2491	GACTCCATCATCTTCTTGGGTGGCTTCCGCTTTTGTCACTCAGAAAGACTGGGTTTGGG	2550
QY	2203	CTAAACACTACCTCCAGGTGGAGTCCAAAGCTTCCAAACATTTTTCATTTTGTGAATCTT	2262
Db	2551	TTAAACAACCGTTCACAGAGCTGAAAGCTTCCAAACGGTTTTCCTTTCTGTGAATCTT	2610
QY	2263	CCCTACTCTGTATCAGAGGTGAAGAAATTTGCTTGGAAATPACTATPACTCAATTAATTTG	2322
Db	2611	CCGTACTCTGTATCAGAGGTGAAGAAATTTGCTTGGAAAGTATCATCTGTCAATTAATTTG	2670
QY	2323	AAAGATGCCACTGAGGTTAAAGGTAAATCTATGAAAAGTGAACAAATTTGATATTTAAATG	2382
Db	2671	AAAGATPACAATTAAGTGTGTGATATCTCATGTAGGAAATGACAGCTTTGATATTTTAAATG	2730
QY	2383	ACTTCAATGAAATTAATATGCCACAGGCCACCAAGACAGACCTTCTGTCTTCCAGTGAAGAT	2442
Db	2731	ACTTCAAAATGACACCAATGGCACCATATPACCGGAAGACTGTTCAGATTTCCAGAGGATAT	2790
QY	2443	GGGGCAACTGTCTTTTCCCTCAGGCGCAACATCTGGGGAATAATCTCATCAAGTCTC	2502
Db	2791	GGGGTCACTCTTGTTTTCCCTCAGGCGCAAGCCACAGCACTTGGAGAGATTTCCATCAAGT	2850
QY	2503	ACAGCTCTTTCACCCACTGCTTCTGTGATCTGTCAACCAAGATGATTTAGTAAAGCTGAA	2562
Db	2851	ACGGCTGCTGCACCCACTGCTGTGATCTGTCAACCAAGCAATGTATGTAAGCTGAA	2910
QY	2563	GGAAATGAAAAATCATATTACAAATCATCTTATTAAGCTTGACTGACAAATAGGCTACAG	2622
Db	2911	GGGAATGAAAAATCCTATTTCCAAATCTGTATTTGATCTGTACAGATAGCAAGTGA	2970
QY	2623	AGTACCCCTGAAAACTTTGAGTTTCTCATTTCCCTCTATPACAGAGACTGGAGTGAAGA	2682
Db	2971	AGTAAAGCAAGCAATCTATAGAGATTTCTTCTCTCTGATPAGGTCAATTTGGAGTGAAGA	3030
QY	2683	GTTCCAGATCACTGCAATTGAGAGATTTCTTGGTCTCTTCATCAATAGGCTTACCTCATYTG	2742
Db	3031	GTTCCAGATCAAGCAATTTGAGAGATATCCTTGGTCTCTCAATCAATAGGCTTATCTTCACTG	3090
QY	2743	ATTGCGATGECTTATAGCTGTGTGTGAACAGAAACATGATPAAATTTTGTCTCCAAATATTAC	2802
Db	3091	ATCCGGAAGCCTTATAGATGTGTGTGAACAGAAACATGATPAAATTTTGTCTCCAAATATTAC	3150
QY	2803	ATTTTGAATTAATCTGACTPAAAAAGAAACAACGACAGATPAAATTTTGAAGAAAAAGCTCTT	2862
Db	3151	ATTTTGAATTAATCTGACTPAAACAGAAACAGCTGAAGTAAATTTTGAAGAAAAAGCCTTT	3210
QY	2863	TCAATTAAGAGCAAGTTTACACAGAGAAACTTCTCTATCAGAGGGAAGATGAGCTTTTC	2922
Db	3211	TCAATTAAGAGCAAGTTTACCAAAAGGAGCTTCTCTATCAGAGGGAAGATGAGCTCTTTC	3270
QY	2923	AGTGTCTTTGGGAATTAATGACCCCTTCTGGAGACCTTGGTGTCAAGCTTTTGTTTTAAAG	2982
Db	3271	AGTGTCTTTTGGGGAATTAATGACCTTCTGGAGAGCACTTGGTGTCAAGCTTTTGTTTTAAAG	3330
QY	2983	TGTTTCTCTGAAGCCGATCTTACATAGATATTGATTCGAATGTGTATCACAGAAACATAC	3042
Db	3331	TGCTTTCTCGAAGCTGATTAATCATATTAATATATGTATTCAGATGTGTATCACAGAAACATAT	3390
QY	3043	ACTTGGCTTAAAGGACATCAGAAATCCAAAGGTGAATTTTGGGATCCAGGAAGGTGATTT	3102
Db	3391	ACTTGGCTTAAATGCAATPAGAAATTTCAATGTGTAAATTTTGGAGCCAGGAAGAGTGAATTT	3450

QY	3103	CAGTGTAGCTTCAAGGGGCGCAATTAAGTCCAGTAACCTTAAGCTTAATTTGTACT	3162
Db	3451	CACAGTGAACCTTCAAGGGGCGCACAAAGGCCCAATACACTTAAGGGCTTAATTTGTACT	3510
QY	3163	TCTCTCCGCGGATATTAGAAGTATCAGGCTTAACCTTAATGTGCAGAAGTCTATCATTTT	3222
Db	3511	TCTGTCTCGGATACAAAAAGTATACAGCTTAATTCAGATGTACAAAGTCAATCAAGTT	3570
QY	3223	TTGAGTCTGAATTCAGTAGAGAAATTTCAGACAATTAACTTAAGCTTTAATACCTTAT	3282
Db	3571	TTGGAAATTTGAATTCAGCAGAGGAATTTGGACAAATTATACCTTAAGCAATTAATTCAT	3630
QY	3283	GCATTGTCAATAGTGGGGAAGTCTTAAGCCGAAGGAAGCTTTGAATATGTGACCTTGAGA	3342
Db	3631	GCCCTGTCAACAGTCCGGAGGCCCTTAAGCAGAGGAGCTTTGAACTTGATGATCAGGA	3690
QY	3343	GCAGAACCAAGAAGGTGGCATCAATTTCTGGGTGTCACTCAGAGTCCAAACTTTCTGACTCC	3402
Db	3651	TCAGAAAAGAAAGACACACACTTTCTGTGTCTATCAGACCTCGCCTCTCTGGGTCC	3750
QY	3403	TGGACCCAGCGCTCCCTGGATATTGAAGTGGACCTATGACCTGTCTCACACTTCTTA	3462
Db	3751	TGGACGCCGCGCTCCGTGTGACATCGAAATCGCTCTTAATGACCTGTGGCGCACACTCTG	3810
QY	3463	CAATTTGACATTTCTGAGGGAATTCCCAATTATAGGTGTGCTTAACACAGCAAAAGAAATAGC	3522
Db	3811	CA---CAATGTGTCTGAGGAATCCCGGTATATGAGTGGCTATCTCAGCAAAAGAAACGC	3867
QY	3523	TTGGGTGGTATTGGATCTACTCAGAGATACCACTGTGCTTTAAGGCTCTGTCTGAATTT	3582
Db	3868	CTGGAGGTTTGTATTCACCTCAGAGATATCTGTTGTGGCCCTGAAGGCAATTATCTGAATTT	3927
QY	3583	GCAGCCCTTAATGAATACAGAAAGACAATAATCCAGTGAACGTGACGGGCGCTTAGCTCA	3642
Db	3928	TCAGCCCTTGTGCAACAGAAATAATACAGATATTCACATGACCGTCAAGGGCCGGCATC	3987
QY	3643	CCAACTCT-----	3651
Db	3988	CCCAATCTCAATACACTTCCGAATTAAGTCCCAAGACCTTTTCTCTTACACAGAAAG	4047
QY	3652	CTTGCTGTGTGACAGCCAAATGGCAGTTAAATTTCCGCAATGGTTTTGATTTGCTATT	3711
Db	4048	CTTCAATGCGCTAGATCCCATACCGTTAAATTTTCTGGCATGTGCTCGGATTTTGCTATT	4107
QY	3712	TGTGAGCTCAATGTTGTATATATATATATGTAAGGCTTCTGGCTCTTCTAGAAAGCAAGATCT	3771
Db	4108	TGCCAGCTTAATGTGACTAATAAGTAAAGGTTCAAGTTCTTTTAAAGACGAGATCT	4167
QY	3772	ATCCAAATCAAGAAGCCTTGATTTAGATGTCTGTAAAGAAATTAAGATGATCTC	3831
Db	4168	ACCGAAATCAAGAAGTTTGTGATTTAGACGTCAATTTGA---ATAATGAGGAAGCAACTT	4224
QY	3832	AATCATGTGATTTGAATGTGTATCAAGCTTTTCGGGCCCGGGTAGAGGTGATGGCT	3891
Db	4225	AGTCACTGAATCTGAATGTGTGACAAAGTCACTTGGGTTTCAGAGAGACAGGATGGTC	4284
QY	3892	CTTAATGAGTAACCTAATTAAGTGGCTTTATGTGTGCTTCAGAAAGCAATTTCTGAGC	3951
Db	4285	CTTAATGAGTAAGCACTTCTCAGTGGCTTTAGTGCACTTTCAGATTCATTTCTCTAGT	4344
QY	3952	GAGACAGTGAAGAAAGTGAATATGATCATGGAAGAACTCAACTTAATTTAGATTTCTGA	4011
Db	4345	GAGACCTGAAGAAAGTGAATATGATTAAGGGAACCTCAACTTTAATTTAGATTTCTGA	4404
QY	4012	AATGAAACCCAGTTTGTGTTAATATTCCTGTCTGTGAAACTTTAAGTTTCAAAATACC	4071
Db	4405	AATGAATCCAGTTTGTGTAAATATTCCTGTGTGAGAGCTAACAAGTTTCAAAATAT	4464
QY	4072	CAAGATGCTTCAAGTGTCCATATGTGATTACTATAGCCAAAGAGACAGGCGGTGAGAGT	4131
Db	4465	CGAGATGCTTCAAGTATCTGTATGTGACTATATTAACCAAGAGACAGGCGGTGAGAGC	4524
QY	4132	TACAACTCTGAAGTGAAGCTGTCTCTCTGTGACCTTTGCAAGTATGTCAGAGGCTGCGT	4191

	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	PUBMED	REFERENCE	AUTHORS	TITLE	JOURNAL	REMARK	COMMENT
D6	4525	TACAACACCAGGTAAGCTGTCCATGGTAACT--CAATTCAGACACCAACTCGAAG	4581																
OY	4192	CCTTGATGANTGAGCTTCAGAGCTGCCCATCATCATTCTTGAAGTCATTTATTTCTGT	4251																
D6	4582	TCACACACGACGAGCCACGACATCCCCTTGCAGCGTTCTTGCAGCTTCTGTTC	4641																
OY	4252	TTCAAGCTTCTGTACTTATGAACTTTGGCTGGA	4287																
D6	4642	TCCTGCTTCTGTACTTGTTCATCATTTGATTAAGA	4677																
RESULT 11	BC052443	BOS2443	5644 bp	mRNA	linear	ROD 19-MAY-2003													
LOCUS	BC052443	Mus musculus cDNA clone MGC:63463 IMAGE:6406841, complete cds.																	
DEFINITION	BC052443																		
ACCESSION	BC052443	.1 GI:30851465																	
VERSION	MGCL																		
KEYWORDS	Mus musculus (house mouse)																		
SOURCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;																		
ORGANISM	Strasbourg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shamen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buatou, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heide, F., Ditichenko, L., Marudina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stepleton, M., Soares, M.B., Donald, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Udell, T.B., Toohy, P., Carninci, P., Prange, C., Raha, S.S., Loggiano, N.A., Peters, G.J., Abramson, R.D., Molloy, S.J., Bosak, S.A., McEwan, P.J., McCrean, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Wolley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulys, S.W., Vallalath, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bonfield, W.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickinson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E., Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A.																		
REFERENCE	1 (bases 1 to 5644)																		
AUTHORS	Strasbourg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shamen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buatou, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heide, F., Ditichenko, L., Marudina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stepleton, M., Soares, M.B., Donald, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Udell, T.B., Toohy, P., Carninci, P., Prange, C., Raha, S.S., Loggiano, N.A., Peters, G.J., Abramson, R.D., Molloy, S.J., Bosak, S.A., McEwan, P.J., McCrean, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Wolley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulys, S.W., Vallalath, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bonfield, W.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickinson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E., Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A.																		
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences																		
JOURNAL	Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)																		
MEDLINE	22388257																		
PUBMED	12477932																		
REFERENCE	2 (bases 1 to 5644)																		
AUTHORS	Strasbourg, R.																		
TITLE	Direct Submission																		
JOURNAL	Submitted (15-May-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA																		
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov																		
COMMENT	Contact: MGC help desk Email: cgabbs@mail.nih.gov Tissue Procurement: Dr. Jim Ivin, University of Iowa cDNA Library Preparation: M. Bento Soares, University of Iowa DNA Sequencing by: The I.M.A.G.E. Consortium (LNL) Thomas L. Casavant. Web site: http://genome.uiowa.edu Contact: bento-soares@uiowa.edu; tom-casavant@uiowa.edu Bonaldo, M.F., Akabogu, I., Bair, T., Bair, J., Crouch, K., Davis, A., Fichter, K., Keppel, C., Kucaba, T., Lebeck, M., Melo, A., Schaefer, K., Scheetz, T., Smith, C., Snit, E., Tack, D., Trout, K., Walters, J., Casavant, T., Soares, M.B.																		

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLT at: http://image.lnl.gov

Series: Plate: Row: Column: 0.

```

source
1. .5644
/organism="Mus musculus"
/mo_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="MGC:63463 IMAGE:6406841"
/issue_type="Brain, enriched mouse brain 12.5dp"
/clone_id="NH_BMAP_F00"
/lab_host="DH10B"
/notes="Vector: pTX-ASC"
105. .4433
/codon_start=1
/product="unknown (protein for MGC:63463)"
/protein_id="AAH52443.1"
/db_xref="GI:30851466"
/translation="MRSRLRLLSAHLCLCAVALAAGSRFLYTPAGIIRPGAVNTIG
VDLEENSPQVLAQVFKIASNKSRIILAEGFHGRKTLVPLPLSDAKIYE
LHNGSENEIVFSNRIQLFPESKISVLLOTBAKPKPOEYERVLITCSDLKPR
TSYDIFDKDPSKNIQMFSPKDKDGLVSTKPSLIDGMSIOVYNDDQYGSF
OULEYVLPKEVTVQITPLYCSLKSQNGSVIATYTGKPVKSLSLTPFLPSFGKK
KNITKSEINGFANFSPDNEMKVMNLKPLTVSEGSYKLVNVPSPGPAEIIATVEE
SLGISEMASTNVEFKQHDYIERFDYVTKLSLNTATVAKYRSRSGNLOTPEEIN
DLVYVTKRKNHESORDQEMOYIOWNTYIPONGIKLEFPMSISGELQKAYER
DGSSTVYHSMFSPSKTYIOLKTRDXYIVGSPFDLMGSPKROLSTWVLSKQGL
VAAKQSSRTFSLTPESMAPKACTIAYIAEDBEIINDLIKLPQVLPENKXLPMS
KPLPSPDKVSLRSITQSDSLVGIADVSLMEKNSRIIMETWHELELVNTEY
LGFEMSPFAVFOBCGLWATDAILRISIDBYDTEYSRFAEENABELVTEYYS
VNNVYAKNPEPTWIMLDAYMGSKIYEPEFVTVDSITSWSAFVISEDGFELTV
PAELOAFQPEPLFLNLPYSVIRGEPELVSIYVLYKDTIKVTLIBESPFILMTS
NDNGTITRKTYOVRPNDGTVIPEIKPHLGEIPITVVASPASADQOTVYKPEE
GIEKSYKSLVLLDITBDSNYSKQDSMRFSPTPTVIGSERVOITADIGLSINGLS
SLIRMPYGCSENNITFAPNIVYILDYVTKQKQLVNLKEKALSTWQGYORELLYGR
DGSFASRQIDSSGTSWLSAFVLRCPLEADYIYDIDDDVLRITVNLAKHGFEPWE
EPGVHISELQGGTSKSPVTLAYIVTSVYKPKQPIVDVDSIKFLPEPERSRISDN
YVTLAISYALSTVSGSPAEKALNLMQMSRKEGDTQWLSGSPALSGSMOBSYDI
AAVALAHLTHVYSEGIPIVMBRLIQONSICGPTVSDTYVALKALSEFALYHKE
DIQITTVGPETPRSIHFRIDSONFLHQEELHLDETVNVSAHSGFALQQLVND
LVGSGSSKRRSTENQGVEDLVIVNNEDDISHLNINQCTSHLGSERTQVNEVNI
NKSASASDSIPLSELTAKVEYNGKMLLDVSNESQFCNIPTVADKYVSNIRDGS
VSWMDYHPRQRAVRSYNTQVKKSSCYLSPTNCKSHTDGATDLSRRSSILVFCSVL
LYPQNH"
BASE COUNT 1556 a 1235 c 1251 g 1602 t
ORIGIN
Query Match 63.2%; Score 2709.2; DB 10; Length 5644;
Best Local Similarity 77.9%; Pred. No. 0;
Matches 3393; Conservative 0; Mismatches 873; Indels 90; Gaps 7;
QY 1 ATGAGAGGCCACCGCTCTCTGACCGCGCCCACTCTCTGCTGTGCACCGCGCGCTG 60
Db 105 ATGCGCAGTGGAGGCTCTCTGAGCGCGCCCACTCTCTGCTGTGCAGCGGCTG 164
QY 61 GCCGTGCTCCCGGCGCTCGTTCTGTGTCACAGCCCCCAGGATCATCAGGCCCGGAGA 120
Db 165 GCC---GCCCCAGGGTCCCGCTTTCTGTGTCACAGCCCCCAGGATCATCAGCCGGAGCA 221
QY 121 AATGTGACTATTTGGGGTGGAGCTTCTGGAACACTGCCCTTCCACAGGTCGATGTAAGCG 180
Db 222 AATGTGACTATTTGGGGTGGAGCTTCTGGAATAATGCCCCCAGGATCTTGTGTAAGGCT 281
QY 181 GAGCTGCTCAAGACAGCATCAAACTCACTGTCTCTGTCTCTGAGAGCAGAGAGTCTTT 240
Db 282 CAGGTGTTCAAGATAGTCTTCAACAATTCAGATTCATCTCTGGAAGCAGAAAGAGTCTTT 341
QY 241 GAAAAAGGCTCTTTAAGACACTTACTCTTCATCACTACTCTCTGAAACAGTGCAGATGAG 300
Db 342 CACGAGAGCCATTCAAGACTCTCTGTTCTTCGCGCACTACTCTGACAGAGGCAGATAG 401
QY 301 ATTATATGAGCTACGCTGTAACCGGACGTAAACCAGAGATGAGATTTTATTTCTTAATATGACC 360
Db 402 ATTATATGAGCTACGATTAACCGGACATCAGAGATATGATTCGATATTTCTTCCACAGAGACA 461
QY 361 CGCTTATCATTTGAGACCAAGAGATATCTGTCTTCAATTCACCAAGACGACCTTATAC 420

```

Db 462 CGTTAACTTTGAGAGAGAGATATCTGCTCCATTCAGAGAGAGCAATTCAC 521  
Qy 421 AAGCCAAAGAGAGAGAGATTTGAGATTTGACATCTTCTGAGATTTGAGCTTAC 480  
Db 522 AAGCCGAGAGAGAGAGATTTGAGATTTGAGATTTGAGATTTGAGATTTGAG 581  
Qy 481 AAAACCTCTTTAAACATTTCTGATTAAGAGAGAGAGAGAGAGAGAGAGAG 540  
Db 582 AGGACATCTGAGAGATTTTATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 641  
Qy 541 TCACAGAGAGAGATTTGAGATTTTCAAAATTTTCAAGATTTTCAAGATTTTCA 600  
Db 642 TCTCAGAGAGAGATTTGAGATTTTCAAAATTTTCAAGATTTTCAAGATTTTCA 701  
Qy 601 CTGAGAGAGATTTGAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAG 660  
Db 702 TTGAGAGAGATTTGAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAG 761  
Qy 761 GTTTCAGAGATTTGAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAG 720  
Db 762 GTTTCAGAGATTTGAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAG 821  
Qy 721 ATGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTT 780  
Db 822 TTGAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAG 881  
Qy 781 AAGAGAGAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCA 840  
Db 882 AAGAGAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAG 941  
Qy 841 ACAAACATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAG 900  
Db 942 ACAAACATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAG 1001  
Qy 901 AATGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGAT 942  
Db 1002 AAGGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGAT 1061  
Qy 943 CTATCTTCCCTGAGAGAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAG 1002  
Db 1062 CCGTCTTCTGAGAGAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGAT 1121  
Qy 1003 TCAAGAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAG 1062  
Db 1122 TCAAGAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAG 1181  
Qy 1063 GATTAATCTGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCA 1122  
Db 1182 GATTAATCTGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCA 1241  
Qy 1123 GCTGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGAT 1182  
Db 1242 TCTGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTT 1301  
Qy 1183 CAGAGAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAG 1242  
Db 1302 CAGAGAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAG 1352  
Qy 1243 GTTTCAGAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAG 1302  
Db 1353 ATTCAGAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAG 1412  
Qy 1303 CTGAGAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAG 1362  
Db 1413 ATTCAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAG 1472  
Qy 1363 GCAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCA 1422  
Db 1473 ACAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCA 1532  
Qy 1423 GAAATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCA 1482

Db 1533 GAAATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCA 1592  
Qy 1483 GAGTTAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGAT 1542  
Db 1593 GAGTTAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGAT 1652  
Qy 1543 ACAATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCA 1602  
Db 1653 AGAATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCA 1712  
Qy 1603 TATATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCA 1662  
Db 1713 TATATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCA 1772  
Qy 1663 TTTTAAATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTT 1722  
Db 1773 TTTTAAATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTT 1832  
Qy 1723 TCTTTAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGAT 1782  
Db 1833 TCTTTAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGAT 1892  
Qy 1783 AGTGTATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTT 1842  
Db 1893 AGTGTATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTT 1952  
Qy 1843 GAATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAG 1902  
Db 1953 GAATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAG 2012  
Qy 1903 GAATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAG 1962  
Db 2013 GAATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAG 2072  
Qy 1963 TATGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCA 2022  
Db 2073 TATGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCA 2132  
Qy 2023 ATTATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAG 2082  
Db 2133 TTTTAAATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTT 2192  
Qy 2083 ATTATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAG 2142  
Db 2193 ATTATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAG 2252  
Qy 2143 GATTTATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCA 2202  
Db 2253 GATTTATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCA 2312  
Qy 2203 CTAAATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAG 2262  
Db 2313 TTTTAAATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTT 2372  
Qy 2263 CCTATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAG 2322  
Db 2373 CCTATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAG 2432  
Qy 2323 AAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAG 2382  
Db 2433 AAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAG 2492  
Qy 2383 ACTTAAATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAG 2442  
Db 2493 ACTTAAATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAG 2552  
Qy 2443 GGGGAAATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAG 2502  
Db 2553 GGGGAAATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAG 2612  
Qy 2503 ACAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAG 2562  
Db 2613 ACAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAG 2672

QY 2563 GGAATAGAAAAATCATATTCACATTCATTATAGACTTGACATAGAGCTACAG 2622  
 DB 2673 GGGATAGAAAAAGTCATATTCGAAATCTGTATTGATCTGACAGATAGCAACGTAGAA 2732  
 QY 2623 AGTACCTGAAAACTTTGAGTTTCTCATTTCTCTCTAATACAGTACGCTGAGAAAGA 2682  
 DB 2733 AGTAGACAGCAATCATAGATTCCTCTCTCTGATACGCTATGGCGCTGAGAAAGA 2792  
 QY 2683 GTTCAGATCACTGCAATGAGAGATGTTCTGTCTCTCATGATGGCTAGCTCATTTG 2742  
 DB 2793 GTTCAGATCACTGCAATGAGAGATGTTCTGTCTCTCATGATGGCTAGCTCATTTG 2852  
 QY 2743 ATTGCGATGCTTATGCTGTGTGTGAGACAGAACATGATTAATTTTGTCTCAATATTTAC 2802  
 DB 2853 ATCCGATGCTTATGATGTGTGTGAGACAGAACATGATTAATTTTGTCTCAATATTTAC 2912  
 QY 2803 ATTTGGATTTATCTGATCTAATAAAGAAACACTGACAGATTAATTTAAGAAAGCTCTT 2862  
 DB 2913 ATTTGGATTTATCTGATCTAATAAAGAAACACTGACAGATTAATTTAAGAAAGCTCTT 2972  
 QY 2863 TCATTTAGAGGCAAGTTACAGAGAGAACTTCTATCAGAGGAGAGATGGCTCTTC 2922  
 DB 2973 TCATTTAGAGGCAAGTTACAGAGAGAACTTCTATCAGAGGAGAGATGGCTCTTC 3032  
 QY 2923 AGTGTCTTTGGGAATTAAGACCTTCTGGAGACACTTGGTGTGAGCTTTTGTTTAAGA 2982  
 DB 3033 AGTGTCTTTGGGACATTTGACTCTCTGGAGACACTTGGTGTGAGCTTTTGTTTAAGA 3092  
 QY 2983 TGTTCCTTGAACCGGATCTTATCATATATTTATCATATGATGTGTACAGAAACATAC 3042  
 DB 3093 TGTTCCTTGAACCGGATCTTATCATATATTTATCATATGATGTGTACAGAAACATAC 3152  
 QY 3043 ACTTGGCTTAAAGCAATCAAGAAATCCAAAGGTGATTTTGGATCCAGAGAGATGAT 3102  
 DB 3153 ACTTGGCTTAAAGCAATCAAGAAATCCAAAGGTGATTTTGGATCCAGAGAGATGAT 3212  
 QY 3103 CATATGAGCTTCAAGTGTGCAATTAAGTCCAGTAAACACTTACAGCTTATTTGTAAT 3162  
 DB 3213 CACAGTGAATCTCAAGTGTGCAATTAAGTCCAGTAAACACTTACAGCTTATTTGTAAT 3272  
 QY 3163 TCTCTCTGGGATATAGAAAGTATCAGCTTAACTTATGATGACAGAGTCTATTCATTT 3222  
 DB 3273 TCTCTCTGGGATATAGAAAGTATCAGCTTAACTTATGATGATGACAGAGTCTATTCATTT 3332  
 QY 3223 TTGAGTGTGATTTCAATGAGAGATTTCAAGCATTTATCTGAGCTTATTAATTTAT 3282  
 DB 3333 TTGAGTGTGATTTCAATGAGAGATTTCAAGCATTTATCTGAGCTTATTAATTTAT 3392  
 QY 3283 GCATTGTATCATGATGGAGAGTCTTAAAGCGAAGAGCTTGAATATGCTGACTTGAGAG 3342  
 DB 3393 GCATTGTATCATGATGGAGAGTCTTAAAGCGAAGAGCTTGAATATGCTGACTTGAGAG 3452  
 QY 3343 GCAGAACAGAGAGTGTGATGCAATTTCTGGTGTATCATGAGAGTCTTGTACTTC 3402  
 DB 3453 TCAGAAAGAGAGAGAGAGACACAGTTCTGTGTATCATGAGAGCTGTCTGTGGTCTC 3512  
 QY 3403 TTGAGGACCACTGCTCTGATATTTGAAGTTGAGAGCTTATGCTGCTCATCTTTCTTA 3462  
 DB 3513 TTGAGGACCACTGCTCTGATATTTGAAGTTGAGAGCTTATGCTGCTCATCTTTCTTA 3572  
 QY 3463 CAATTTCAGACTTTCAGAGGAAATCCCAATTTATGAGTGTGCTAAGAGCAAGAAATAGC 3522  
 DB 3573 CA---CAATGTCTTGAAGGAAATCCCGTTATGAGTGTGCTAATCAAGAAAGAAACAGC 3629  
 QY 3523 TTGGGTGTGTTTGCATCTACTAGAGATACCATGTGTGCTTTAAAGCTTGTCTGAATTT 3582  
 DB 3630 CTGGAGAGGTTTGTATCTACCTAGAGATATGTTGTGGCCCTGAGAGCAATTTACTGAATTT 3689  
 QY 3583 GGAAGCCCTATGATATACAGAAAGAGCAATATTCAGATGACGCTGACGGGCTTACGCTCA 3642  
 DB 3690 TCAGCCCTTGTGCAACAGAAATATCAGATATTCAGATGACGCTGACAGGGCCCGGCAATC 3749

QY 3643 CCAAGTCT----- 3651  
 DB 3750 CCCAGATCATATACACTTCCGATTTGACTCCAGAAACCTTTCTTCTTCAACAGAGAG 3809  
 QY 3652 CTGTGTGTGTACAGCCAAATGGCAATTAATTTCCGAAATGTGTTGATTTGCTATT 3711  
 DB 3810 CTTCATGGCTATGATCCCATACCGTTAATGTTCTGAGCATGGCTCGGATTTGCTATT 3869  
 QY 3712 TGTGCTCATGATGTTGATATATATATGAGAGGCTTCTGGGTCTTCTAGAAAGCAAGATCT 3771  
 DB 3870 TGTGCTCATGATGTTGATATATATGAGAGGCTTCTGGGTCTTCTAGAAAGCAAGATCT 3929  
 QY 3772 ATCCAAATCAAGAGACCTTGTATTTAGATGTTGCTGTAAAGAAATTAAGATATCTC 3831  
 DB 3930 ACCGAAATCAAGAGATTTTGTATTTAGATGATGATGATGATGATGATGATGATGAT 3986  
 QY 3832 AATCATGTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3891  
 DB 3987 AGTACCTGAAATCTGAATGTGTGTGACAGATCACTTGGGTGTGACAGAGAGAGCATGGTC 4046  
 QY 3892 CTATGAGAGTTAATCTTATTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3951  
 DB 4047 CTATGAGAGTTAATCTTATTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4106  
 QY 3952 GAGACAGTGAAGAAAGTGAATATGATCATGAGAAACCTCACTTATTTAGATTTCTGTA 4011  
 DB 4107 GAGACCTGTAAGAAAGTGAATATGATCATGAGAAACCTCACTTATTTAGATTTCTGTA 4166  
 QY 4012 AATGAACCCAGTTTGTGTATATTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4071  
 DB 4167 AATGAATCCAGTTTGTGTATATTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4226  
 QY 4072 CAAGATGCTGAGT 4131  
 DB 4227 CGAGATGCTGAGT 4286  
 QY 4132 TACAACCTGAGTGAAGAGT 4191  
 DB 4287 TACAACCTGAGTGAAGAGT 4343  
 QY 4192 CTTGT 4251  
 DB 4344 TCACACAGGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAG 4403  
 QY 4252 TTCAAGCTTCTGTATTTATGAACTTTGGCTGTGA 4287  
 DB 4404 TCCGTCTTGT 4439

RESULT 12  
 HSM805581  
 LOCUS Homo sapiens mRNA; cDNA DKFZp762L1111 (from clone DKFZp762L1111).  
 DEFINITION  
 AL834478  
 ACCESSION  
 VERSION AL834478.1 GI:21740252  
 KEYWORDS  
 SOURCE  
 ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE  
 1 (bases 1 to 2938)  
 TITLE Direct Submission  
 AUTHORS Blum, H., Baurasche, S., Mewes, H.W., Weil, B. and Wiemann, S.  
 JOURNAL Submitted (09-JUL-2002) 1, D-85764 Neuberger, GERMANY  
 COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
 Research Center (DKFZ), Email s.wiemann@dkfz-heidelberg.de;  
 sequenced by IMU (Ludwig Maximilians University,  
 Munich/Germany) within the cDNA sequencing consortium of the German  
 Genome Project.  
 This clone (DKFZp762L1111) is available at the RZPD in Berlin.  
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further  
 information about the clone and the sequencing project is available

at <http://mips.gsf.de/proj/cDNA/>.  
 FEATURES  
 source  
 Location/Qualifiers  
 1. 2938  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /map="15"  
 /clone="DKFZp762L1111"  
 /tissue\_type="melanoma (Melo cell line)"  
 /clone\_lib="762 (synonym: hmel2). Vector pSPORT1; host DH10B; sites NotI + SalI"  
 /dev\_stage="adult"  
 1. 2938  
 /gene="DKFZp762L1111"  
 /gene="DKFZp762L1111"  
 /note="similarity to alpha-2-macroglobulin (Limulus sp.)"  
 /codon\_start=1  
 /product="hypothetical protein"  
 /protein\_id="CAD39137.1"  
 /db\_xref="gi:21740253"  
 /translation="VDSKSVNLMNASNDITMENVVHELELYNTGYILAMFNNSFAVPOE  
 CGLMVLTDANITKDYIDGYDNAEYARPMENEGHVIDIHDSLSPPHYRKHPEET  
 WMLDWMGSKTYQEFVETVDSITSVATGFISEDLGGLTTPVELDAPQPF1E  
 LNLPSYVIRGEFALBITIENYLDATKEVRIEKSDFILMTSSINATGHQOTLL  
 VPSEBDGATVLPFIPITLGEIPITVATSPASDAITOMILVAKGIEKSYOSILDL  
 LDNRLOSTLKTLSFSPFNPVTGSEVQITAGIDVAGPSINGLASLIEMPGSEGN  
 MINPAPNITLDVLTKKOLTDNKEKALSPMGOGYOBELLXOEDSPSAPRYVDS  
 GSTSLAFVLRCLFADRPYIDIDQVILHRTYTLKQKXNGEWDGRVTHSLQSG  
 NKSPVTLAVYIVSLGIRYOPNIDVESIHFLSEFSRGISDNYTLALITYALSQV  
 GSPKAKALNMLTWRAEOEGMOPVWSESLLDSQPSRIDIEVAAYALSHLQPO  
 TSEGIPIRMWLSRQNSIGFASOTDTVALKALSEPAALMTERTNIQYLVGSPSP  
 SPVKEFLDTHNRLLQTAELAVQPTAVNISANGEPALCOLAVVNVKXSGSRBR  
 SIONOEAFDLDIVAKENKDLNHYDLNVCISFSGRSGMALMEVNLSPMPVSEAI  
 SLSTTKVKEIDHCKNLMLYLDVSUETQPCNIPVRKFSNTODASVSYDIYEPR  
 QAVRSYNSVRLSSCDLCSDVQGRPEDGASGSHHSVLFICFELLYPMEIWL"  
 2843. 2848  
 /gene="DKFZp762L1111"  
 /gene="DKFZp762L1111"  
 /gene="DKFZp762L1111"  
 /gene="DKFZp762L1111"  
 BASE COUNT 887 a 529 c 620 g 902 t  
 ORIGIN  
 Query Match 57.1%; Score 2448.2; DB 9; Length 2938;  
 Best Local Similarity 97.9%; Pred. No. 0;  
 Matches 2511; Conservative 0; Mismatches 3; Indels 51; Gaps 1;  
 polyA\_signal  
 polyA\_site  
 1774 GTTGACAAAAGTGAATCTGATGAATGCTTAATGATATTACATGAAAATGTGCTC 1833  
 Db 1 GTTGACAAAAGTGAATCTGATGAATGCTTAATGATATTACATGAAAATGTGCTC 60  
 Qy 1834 CATGAGTTGGAACCTTTATTAACAAGATATTATTAAGCAATGTTCAATGATCTTTTGA 1893  
 Db 61 CATGAGTTGGAACCTTTATTAACAAGATATTATTAAGCAATGTTCAATGATCTTTTGA 120  
 Qy 1894 GTCTTCAGAAATGGAAGCTCTGGATTTGACAGATGCAAACTCAGAGAGATTTATTT 1953  
 Db 121 GTCTTCAGAAATGGAAGCTCTGGATTTGACAGATGCAAACTCAGAGAGATTTATTT 180  
 Qy 1954 GATGCTTTATGACAAATGAGAAATATGCTGAGAGTTTATGAGAGAAATGAGAGCAT 2013  
 Db 181 GATGCTTTATGACAAATGAGAAATATGCTGAGAGTTTATGAGAGAAATGAGAGCAT 240  
 Qy 2014 ATTGTAGATATTCAATGATCTTTCTTTGGGTGACAGTCCACATGTCGGAAGCATTTTCCA 2073  
 Db 241 ATTGTAGATATTCAATGATCTTTCTTTGGGTGACAGTCCACATGTCGGAAGCATTTTCCA 300  
 Qy 2074 GAGACTTGAGTTGGCTGAGACCAACATGAGTTTACAGAGTTTACAGAGTTTGAAGTA 2133  
 Db 301 GAGACTTGAGTTGGCTGAGACCAACATGAGTTTACAGAGTTTACAGAGTTTGAAGTA 360  
 Qy 2134 ACTGTACCTGATCTATCACTCTTGGGTGCTACTGTTTGTGATCTGAGAGCTG 2193

Db 361 ACTGTACCTGATCTATCACTCTTGGGTGCTACTGTTTGTGATCTGAGAGCTG 420  
 Qy 2194 GGTCTTGAGTAACTAACAATACTATCCAGTGAAGCTCCAGGCTTCCACCAATTTTCAATTTT 2253  
 Db 421 GGTCTTGAGTAACTAACAATACTATCCAGTGAAGCTCCAGGCTTCCACCAATTTTCAATTTT 480  
 Qy 2254 TTGAATCTTCCCTACTCTGTTTATCAGAGGTGAAGAAATTTGCTTTGAAATTAATATATTC 2313  
 Db 481 TTGAATCTTCCCTACTCTGTTTATCAGAGGTGAAGAAATTTGCTTTGAAATTAATATATTC 540  
 Qy 2314 AATTATTGGAAGAATGCACTGAGGTTAAGCTATATCATTTGAGAAAGTGCAAAATTTGAT 2373  
 Db 541 AATTATTGGAAGAATGCACTGAGGTTAAGCTATATCATTTGAGAAAGTGCAAAATTTGAT 600  
 Qy 2374 AATTCTAATGACTTCAAGTGAATAATTAATGACAGGCGCACAGAGACCTTCTGTGTTCCC 2433  
 Db 601 AATTCTAATGACTTCAAGTGAATAATTAATGACAGGCGCACAGAGACCTTCTGTGTTCCC 660  
 Qy 2434 AGTGAAGATGGGCAACTGTTCTTTTCCCATGAGGCAACACATCTGGAGAAATTCCT 2493  
 Db 661 AGTGAAGATGGGCAACTGTTCTTTTCCCATGAGGCAACACATCTGGAGAAATTCCT 720  
 Qy 2494 ATCAGACTCAGACTCTTTCAACCCAGCTGCTGATGCTGTACCCAGATGATTTTATGTA 2553  
 Db 721 ATCAGACTCAGACTCTTTCAACCCAGCTGCTGATGCTGTACCCAGATGATTTTATGTA 780  
 Qy 2554 AAGGCTGAAGAAATGAGAAATATCATATTCAATCCATCTTATTAAGACTGATGACAAAT 2613  
 Db 781 AAGGCTGAAGAAATGAGAAATATCATATTCAATCCATCTTATTAAGACTGATGACAAAT 840  
 Qy 2614 AAGCTACAGAGTACCTGAAACCTTTGAGTTTCTCATTTCTCTTAATACAGTACTGAC 2673  
 Db 841 AAGCTACAGAGTACCTGAAACCTTTGAGTTTCTCATTTCTCTTAATACAGTACTGAC 900  
 Qy 2674 AGTGAAGAAGTCAATCATCTGCAATTTGAGAGATTTCTGTCTTCCATCAATAGCTTTA 2733  
 Db 901 AGTGAAGAAGTCAATCATCTGCAATTTGAGAGATTTCTGTCTTCCATCAATAGCTTTA 960  
 Qy 2734 GGCCTCATTTGAGATGCTTATGAGCTTATGAGCTGTGTGAGACAGAAACATGATTAATTTGCTCCA 2793  
 Db 961 GGCCTCATTTGAGATGCTTATGAGCTTATGAGCTGTGTGAGACAGAAACATGATTAATTTGCTCCA 1020  
 Qy 2794 AATATTATCACTTTTGTGATTTATCTGACTAAAGAAACCACTGACAGATTAATTTGAAAGAA 2853  
 Db 1021 AATATTATCACTTTTGTGATTTATCTGACTAAAGAAACCACTGACAGATTAATTTGAAAGAA 1080  
 Qy 2854 AAGGCTCTTCAATTTATGAGGCAAGTTTACAGAGAGAACTTCTATTCAGAGGAGAAAT 2913  
 Db 1081 AAGGCTCTTCAATTTATGAGGCAAGTTTACAGAGAGAACTTCTATTCAGAGGAGAAAT 1140  
 Qy 2914 GGCCTTTCAGTCTTTTGGGAATTTATGACCTTCTGGGAGCACTTGGTGTGAGCTTTT 2973  
 Db 1141 GGCCTTTCAGTCTTTTGGGAATTTATGACCTTCTGGGAGCACTTGGTGTGAGCTTTT 1200  
 Qy 2974 GTTTTAAGATGTTTCTTTGGAAGCGATCTTACATAGATATTGATCAGAAATGTGTTACAC 3033  
 Db 1201 GTTTTAAGATGTTTCTTTGGAAGCGATCTTACATAGATATTGATCAGAAATGTGTTACAC 1260  
 Qy 3034 AGAACAATCACTTGGCTTAAAGACATCAGAAATCCACGCTGAATTTTGGATCCAGGA 3093  
 Db 1261 AGAACAATCACTTGGCTTAAAGACATCAGAAATCCACGCTGAATTTTGGATCCAGGA 1320  
 Qy 3094 AGAGTATTCATAGTGAAGCTTCAAGGTGCAATTAAGTCCAGTACCTTACAGGCTAT 3153  
 Db 1321 AGAGTATTCATAGTGAAGCTTCAAGGTGCAATTAAGTCCAGTACCTTACAGGCTAT 1380  
 Qy 3154 ATTGTAACTCTCTCTGAGATATAGAAATATCAGCTTAACATTTGATGTCAGAGCTCT 3213  
 Db 1381 ATTGTAACTCTCTCTGAGATATAGAAATATCAGCTTAACATTTGATGTCAGAGAGTCT 1440  
 Qy 3214 ATCCATTTTGGAGTGTGAATTCAGTAGAGAAATTTACAGCAATTTATCTAGCCCTT 3273  
 Db 1441 ATCCATTTTGGAGTGTGAATTCAGTAGAGAAATTTACAGCAATTTATCTAGCCCTT 1500



```

QY 3274 ATAACTTATGCAATGTCATGAGTGGAGAGCTCTAAAGCAAGAGAGCTTTGAATATGCTG 3333
    |||
    |||
    |||
Db 1501 ATAACTTATGCAATGTCATGAGTGGAGAGCTCTAAAGCAAGAGAGCTTTGAATATGCTG 1560
    |||
    |||
    |||
QY 3334 ACTTGAGAGGAGCAACAAGAGTGGCAATTTGGTGGTGCATCAAGATCCAACTT 3393
    |||
    |||
    |||
Db 1561 ACTTGAGAGGAGCAACAAGAGTGGCAATTTGGTGGTGCATCAAGATCCAACTT 1620
    |||
    |||
    |||
QY 3394 TCTGACTCTGGCAGCAGCTCCCTGGATATTAAGTTGACGCTATGCACTGCTCTCA 3453
    |||
    |||
    |||
Db 1621 TCTGACTCTGGCAGCAGCTCCCTGGATATTAAGTTGACGCTATGCACTGCTCTCA 1680
    |||
    |||
    |||
QY 3454 CACTTCTTAACAATTTGAGACTTCTGAGGAGATCCCAATTATGAGGTGGCTAAAGCAGCA 3513
    |||
    |||
    |||
Db 1681 CACTTCTTAACAATTTGAGACTTCTGAGGAGATCCCAATTATGAGGTGGCTAAAGCAGCA 1740
    |||
    |||
    |||
QY 3514 AAGAAATAGTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 3573
    |||
    |||
    |||
Db 1741 AAGAAATAGTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1800
    |||
    |||
    |||
QY 3574 TCTGAATTTGCAAGCCCTTAATGAATACAGAAAGGACAAATATCCAAATGACCGTGAAGG 3633
    |||
    |||
    |||
Db 1801 TCTGAATTTGCAAGCCCTTAATGAATACAGAAAGGACAAATATCCAAATGACCGTGAAGG 1860
    |||
    |||
    |||
QY 3634 CCTAGCTCAACAAGTCTT----- 3651
    |||
    |||
    |||
Db 1861 CCTAGCTCAACAAGTCTTGAAGTTTGTGATGACACACACACCGCTTACTCTCTGAG 1920
    |||
    |||
    |||
QY 3652 -----CTTGCTGTGGTACACCCCAATGAGCAATTAATTTCCGCAATGGTTTGGCA 3702
    |||
    |||
    |||
Db 1921 ACAGCAGAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1980
    |||
    |||
    |||
QY 3703 TTGTGATTTGTGAGTCAATGTTGTATTAATGTAAGAGCTTGTGGTCTTCTAGAAGA 3762
    |||
    |||
    |||
Db 1981 TTGTGATTTGTGAGTCAATGTTGTATTAATGTAAGAGCTTGTGGTCTTCTAGAAGA 2040
    |||
    |||
    |||
QY 3763 CGAAGATCTATCCAAATCAAGAGGCTTTGATTTAGATGTTGCTGTAAAGAAATAAA 3822
    |||
    |||
    |||
Db 2041 CGAAGATCTATCCAAATCAAGAGGCTTTGATTTAGATGTTGCTGTAAAGAAATAAA 2100
    |||
    |||
    |||
QY 3823 GATGATCTCAATCATGATGATTTGAATGTGTGTGACAGGCTTTTGGGCGCGGTAGAGT 3882
    |||
    |||
    |||
Db 2101 GATGATCTCAATCATGATGATTTGAATGTGTGTGACAGGCTTTTGGGCGCGGTAGAGT 2160
    |||
    |||
    |||
QY 3883 GGCAATGCTCTTATGGAATTAACCTATTAAGTGGCTTTATGCTTCAAGAGCAAT 3942
    |||
    |||
    |||
Db 2161 GGCAATGCTCTTATGGAATTAACCTATTAAGTGGCTTTATGCTTCAAGAGCAAT 2220
    |||
    |||
    |||
QY 3943 TCTCTGAGCGAGACAGTGAAGAAAGTGAATATGATCATGAAAACTCAACCTCTATTTA 4002
    |||
    |||
    |||
Db 2221 TCTCTGAGCGAGACAGTGAAGAAAGTGAATATGATCATGAAAACTCAACCTCTATTTA 2280
    |||
    |||
    |||
QY 4003 GATTTCTTAATGAAAACCAAGTTTGTGTAATATTCCTGCTGTGAAACTTTAAAGTT 4062
    |||
    |||
    |||
Db 2281 GATTTCTTAATGAAAACCAAGTTTGTGTAATATTCCTGCTGTGAAACTTTAAAGTT 2340
    |||
    |||
    |||
QY 4063 TCAATATCCCAAGATGCTTCAAGTTCATAGTGAATTAATTAAGAGCAAGAGAGAGGCGG 4122
    |||
    |||
    |||
Db 2341 TCAATATCCCAAGATGCTTCAAGTTCATAGTGAATTAATTAAGAGCAAGAGAGAGGCGG 2400
    |||
    |||
    |||
QY 4123 GTGAGAAGTAACTCTGAAGTGAAGTGTCTCTCTGTGACCTTTGAGAGTATGTCGAG 4182
    |||
    |||
    |||
Db 2401 GTGAGAAGTAACTCTGAAGTGAAGTGTCTCTCTGTGACCTTTGAGAGTATGTCGAG 2460
    |||
    |||
    |||
QY 4183 GGCTGCGCTCTTGTGAGAGTGAAGTTCAGGCTCCCAATCATCTTCACTCATTTT 4242
    |||
    |||
    |||
Db 2461 GGCTGCGCTCTTGTGAGAGTGAAGTTCAGGCTCCCAATCATCTTCACTCATTTT 2520
    |||
    |||
    |||
QY 4243 ATTTTCTGTTTGAAGTCTTGTACTTATGAAGCTTTGGCTGTGA 4287
    |||
    |||
    |||
Db 2521 ATTTTCTGTTTGAAGTCTTGTACTTATGAAGCTTTGGCTGTGA 2565
    |||
    |||
    |||

```

```

RESULT 13
AK095888
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
Oshima,A., Takahashi-Fujii,A., Tanase,T., Imose,N., Takeuchi,K.,
Arita,M., Mutsaers,K., Yuuki,H., Hara,H., Sugiyama,T., Irie,R.,
Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y.,
Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H.,
Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K.,
Wagatsuma,M., Murakawa,K., Kanehori,K., Sugiyama,A., Kawakami,B.,
Suzuki,Y., Sugano,S., Nagahara,K., Masuno,Y., Nagai,K. and
Isogai,T.
TITLE
NEDO human cDNA sequencing project
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 2273)
AUTHORS
Isogai,T. and Yamamoto,J.
TITLE
Direct Submission
JOURNAL
Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team) ; 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail:genominfo@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB) ; cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
FEATURES
source
1..2273
location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HCHON2006459"
/cell_type="chondrocytes (HC)"
/clone_1id="HCHON2"
/notes="cloning vector: pME18FLJ3-primary culture,
chondrocytes"
66..2063
/notes="unnamed protein product"
/codon_start=1
/protein_id="BAC04642.1"
/db_xref="GI:21755237"
/translation="MQGPPPLTLAAHLCTVTAALAVADGPRELVTAPGIIIRPGNVTI
GVELLEHPCSOVTVKAEILLKASNLTVLAEVGEFGSKFTLLPLPLNSADEIY
ELRVGTRODELIFSNSTRILSPFETRIKISVFIOTDKALYKPOEVEFRIVTLPSDEKPY
KTSNILLIKDKSNLILQOMLSQOSDLGIVISKTFQSLSPHIIIGDMSIIOVANDQTYOS
FOVSEYVLPKREYTIQTPLYCSMNSKHLNGITTAQTYTGKPKGQVITFLPLSMWK
KNNTITKRLINGSANFSFNDSEMKVMSNGLSFYLDLSSGPELIIITVYESVTGI
SRNSVTNVPKQHDYILIFPDYTYLAKSLNFTATVAKTRADGNOLTEBRANNVIT
VTORAVTYWESGNSGNQMEKAVOKINTVQSGTFKLEFPLIBESSEIOLKAYFLGS
KSSMAVHSLFKPSKTYIOLKTRDNIKVGSPFELVSGNKRLEKLSYVAVRGDLVA
VGKQNTMSESLTPNSWTPKACVIVYIEDGEITISDLIKIPIQVLFNRIKILVSKY
KAERSEKYSLSISVTPDSIYGIYAVDVSVMANASNTIMENVAHELELYNTGYLIG
MFANSFAYFOEGGLWVLTDANITKQYIDGVIDNTLFGTQDAL"
BASE COUNT
712 a 431 c 462 g 668 t
ORIGIN
Query Match
Best Local Similarity 99.9%; Score 1969.4; DB 9; Length 2273;
Matches 1970; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGACGGGCCCAACCGCTCTGACCGCCGCAACTCTCTGTGCTGACGCGCGCGCTG 60
    |||
    |||
    |||

```

Dh 66 ATGACGAGCCCAACCGCTCTGACCGCCGCACTCTCTGCTGTGTGACCGCGCGCTG 125  
Qy 61 GCCGCGCTCCCGGCGCTCGGTTTCTGTGTGACAGCCCGAGGATCATCAGGCCCGGAGGA 120  
Dh 126 GCCGTGGCTCCCGGCGCTCGGTTTCTGTGTGACAGCCCGAGGATCATCAGGCCCGGAGGA 185  
Qy 121 AATGTGACTATTTGGGGTGGAGCTTCTGTGAACACTGCCCTTGACAGGTGACTGTGAAGCG 180  
Dh 186 AATGTGACTATTTGGGGTGGAGCTTCTGTGAACACTGCCCTTGACAGGTGACTGTGAAGCG 245  
Qy 181 GAGCTGCTCAAGACAGCATCAAACTCATCTCTCTGTCTGTGAAGACGAAAGAGTCTTT 240  
Dh 246 GAGCTGCTCAAGACAGCATCAAACTCATCTCTGTCTGTGAAGACGAAAGAGTCTTT 305  
Qy 241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCATCAGCTTGAAGAGAGTGAAG 300  
Dh 306 GAAAAAGGCTCTTTTAAGACACTTACTCTTCATCAGCTTGAAGAGAGTGAAG 365  
Qy 301 ATTATGAGCTACGCTGTAAACCGACGTACCCAGAGTGAATTTTATCTTAATAGTACC 360  
Dh 366 ATTATGAGCTACGCTGTAAACCGACGTACCCAGAGTGAATTTTATCTTAATAGTACC 425  
Qy 361 CGCTTATCATTTGAGACCAAGAAATATCTGTCTTCAATCAACAGACAGGCTTATAC 420  
Dh 426 CGCTTATCATTTGAGACCAAGAAATATCTGTCTTCAATCAACAGACAGGCTTATAC 485  
Qy 421 AAGCAAGCAAGAAAGTGAAGTTGCACTGTGTACCTCTGAGATTTTAAGCTTAC 480  
Dh 486 AAGCAAGCAAGAAAGTGAAGTTGCACTGTGTACCTCTGAGATTTTAAGCTTAC 545  
Qy 481 AAAACCTCTTTAAACATTTCTAATTAAGAACCCCAATCAATTTGATCCAAAGTGTG 540  
Dh 546 AAAACCTCTTTAAACATTTCTAATTAAGAACCCCAATCAATTTGATCCAAAGTGTG 605  
Qy 541 TCACAAACAAGTATCTTGAGAGTCATTTCCAAAATTTCAGCTATCTTCCATCCATA 600  
Dh 606 TCACAAACAAGTATCTTGAGAGTCATTTCCAAAATTTCAGCTATCTTCCATCCATA 665  
Qy 601 CTGTGTGCTGTATCTCAAGTTCAGGTGAAGTGAACGACATTTATCATCTTTCAG 660  
Dh 666 CTGTGTGCTGTATCTCAAGTTCAGGTGAAGTGAACGACATTTATCATCTTTCAG 725  
Qy 661 GTTTCAGAAATATGATTAACAAAATTTGAAGTGACTTTGACAGACCACTTAATTTGTTCT 720  
Dh 726 GTTTCAGAAATATGATTAACAAAATTTGAAGTGACTTTGACAGACCACTTAATTTGTTCT 785  
Qy 721 ATGAATTCGAACATTTAATGATGACATCAACGCAAGTATACATATGAGGACCGCTG 780  
Dh 786 ATGAATTCGAACATTTAATGATGACATCAACGCAAGTATACATATGAGGACCGCTG 845  
Qy 781 AAAAGAGAGCTTAAGGCTTACATTTTACCTTATCTTTGGGAAAGAAATATTT 840  
Dh 846 AAAAGAGAGCTTAAGGCTTACATTTTACCTTATCTTTGGGAAAGAAATATTT 905  
Qy 841 ACAAAAACATTTAAGATTAATGATGACATCACTTCTTTAATGATGAAGATGAAA 900  
Dh 906 ACAAAAACATTTAAGATTAATGATGACATCACTTCTTTAATGATGAAGATGAAA 965  
Qy 901 AATGTATGATTCCTCAATATGACCTTTCTGAATACCTGATCTATCTTCCCTGACCA 960  
Dh 966 AATGTATGATTCCTCAATATGACCTTTCTGAATACCTGATCTATCTTCCCTGACCA 1025  
Qy 961 GTAAATATTTTAACACAGTGAAGATCACTTACAGTATTTTAAAGAAAGTGAAGACT 1020  
Dh 1026 GTAAATATTTTAACACAGTGAAGATCACTTACAGTATTTTAAAGAAAGTGAAGACT 1085  
Qy 1021 AATGTATTCCTCAAGCAACATGATTAACATCACTTGAATTTTGAATTAATCACTGTCTG 1080  
Dh 1086 AATGTATTCCTCAAGCAACATGATTAACATCACTTGAATTTTGAATTAATCACTGTCTG 1145  
Qy 1081 AAGCATCTCTCAACTTACAGCACTGTGTGAAGGTAATCTGTGTGATGGCAACCACTG 1140  
Dh 1146 AAGCATCTCTCAACTTACAGCACTGTGTGAAGGTAATCTGTGTGATGGCAACCACTG 1205

Qy 1141 ACTTTGAAGAAAGAAATTAATGTAGTATTAACAGTGAACAGAGAAATTAATCTAG 1200  
Dh 1206 ACTTTGAAGAAAGAAATTAATGTAGTATTAACAGTGAACAGAGAAATTAATCTAG 1265  
Qy 1201 TACTGAGCGGATCTTAACAGTGAATCAGAAAATGGAAGCTGTTCAGAAAATTAAT 1260  
Dh 1266 TACTGAGCGGATCTTAACAGTGAATCAGAAAATGGAAGCTGTTCAGAAAATTAAT 1325  
Qy 1261 ACTGTCCCAAGTGAACCTTTTAAGTGAATTTCCATCTCGAGGATTCAGTGAAG 1320  
Dh 1326 ACTGTCCCAAGTGAACCTTTTAAGTGAATTTCCATCTCGAGGATTCAGTGAAG 1385  
Qy 1321 CTACAGTTGAAGGCTTATTTCTTGTGTATTAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1380  
Dh 1386 CTACAGTTGAAGGCTTATTTCTTGTGTATTAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1445  
Qy 1381 AAGTCTCTAGTGAACATCATCACTTAAACAGAGATGAAGAAATTAAGTGAAG 1440  
Dh 1446 AAGTCTCTAGTGAACATCATCACTTAAACAGAGATGAAGAAATTAAGTGAAG 1505  
Qy 1441 TCGCTTTTGAAGTGTGTGTAGTGAAGCAAAAGTGAAGGATTAAGTGAAGTGAAG 1500  
Dh 1506 TCGCTTTTGAAGTGTGTGTAGTGAAGCAAAAGTGAAGGATTAAGTGAAGTGAAGTGAAG 1565  
Qy 1501 GTATCCAGGGGACAGTTGTGCTGTAGGAAACAAATTCACAAATGTTCTCTTTTACA 1560  
Dh 1566 GTATCCAGGGGACAGTTGTGCTGTAGGAAACAAATTCACAAATGTTCTCTTTTACA 1625  
Qy 1561 CCAAGAAATTTCTGAGCTCCAAAAGCTGTGTATTTGTATTAATTAAGTGAAGTGAAG 1620  
Dh 1626 CCAAGAAATTTCTGAGCTCCAAAAGCTGTGTATTTGTATTAATTAAGTGAAGTGAAG 1685  
Qy 1621 GAAATTAATGAATGTCTTAAATTTCCCTGTGAGCTGTCTTTTAAATTAAGTGAAG 1680  
Dh 1686 GAAATTAATGAATGTCTTAAATTTCCCTGTGAGCTGTCTTTTAAATTAAGTGAAG 1745  
Qy 1681 CTATATTTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1740  
Dh 1746 CTATATTTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1805  
Qy 1741 ACAAGCTGAGTCCATATGTTGGAGTTGAGCTGTGTGAACAAAGTGAATTCGATCAAT 1800  
Dh 1806 ACAAGCTGAGTCCATATGTTGGAGTTGAGCTGTGTGAACAAAGTGAATTCGATCAAT 1865  
Qy 1801 GCCTCTTAATGATTAATCAATGGAATATGTGCTCAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1860  
Dh 1866 GCCTCTTAATGATTAATCAATGGAATATGTGCTCAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1925  
Qy 1861 TATTAATTAAGGATGATCATGAATTTCTTTGAGTCTTTGAGAAATGAGTCTGAGGTA 1920  
Dh 1926 TATTAATTAAGGATGATCATGAATTTCTTTGAGTCTTTGAGAAATGAGTCTGAGGTA 1985  
Qy 1921 TTGACAGATCAAACTCAGAGAGATTAATGATGATGATGATGATGATGATGATGAT 1971  
Dh 1986 TTGACAGATCAAACTCAGAGAGATTAATGATGATGATGATGATGATGATGATGAT 2036

RESULT 14  
AX083146  
LOCUS 1300 bp DNA linear PAT 28-FEB-2001  
DEFINITION Sequence 49 from Patent WO0110903.  
AX083146  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
AUTHORS  
TITLE  
1 Yue,H., Lai,P., Tang,Y.T., Bandman,O., Baughn,M.R., Azimzai,Y.,  
Lu,D.A. and Yang,J.  
Proteases and protease inhibitors



JOURNAL Patent: WO 0110903-A 49 15-FEB-2001;  
 Incyte Genomics, Inc. (US)  
 Location/Qualifiers  
 source  
 1.1300  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /note="Incyte ID No: 3658034CB1"

BASE COUNT 353 a 239 c 289 g 419 t

ORIGIN

Query Match 17.0%; Score 730.8; DB 6; Length 1300;  
 Best Local Similarity 93.6%; Pred. No. 1.3e-148;  
 Matches 805; Conservative 0; Mismatches 2; Indels 53; Gaps 2;

3479 AGGGAATCCCAATATATAGAGTGGCTAAGAGGCAAGAAATATGCTGGGAGTTTGGCAT 3538  
 1 AGGGAATCCCAATATATAGAGGCGGCTAAGAGGCAAGAAATATGCTGGGAGTTTGGCAT 60

3539 CTACTCAGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGACGCCCTAATGATA 3598  
 61 CTACTCAGATACCACTGTGGCTTTAAAGG--CTGTCTGAATTTGACGCCCTAATGATA 118

3599 CAGAAAGCAAAATATCCAAAGTGAACCGTGAACGGGCTTACCTCAGTCTCT----- 3651  
 119 CAGAAAGCAAAATATCCAAAGTGAACCGTGAACGGGCTTACCTCAGTCTCTCTGTAAGT 178

3652 -----CTTGCTGTGTACAGC 3667

179 TTCTGATTGACACACAAACCGCTTACCTCTTACAGACAGAGCTTGTGTGTACAGC 238

3668 CAATGCGCATTAATATTTCCGCAATATGTTTGGATTGCTATTGTGACGCTCAATGTTG 3727  
 239 CAATGCGCATTAATATTTCCGCAATATGTTTGGATTGCTATTGTGACGCTCAATGTTG 298

3728 TATATATGTGAAGGCTTGTGGCTTTCTAGAAAGCAAGATCTATCCAAATCAAGAG 3787  
 299 TATATATGTGAAGGCTTGTGGCTTTCTAGAAAGCAAGATCTATCCAAATCAAGAG 358

3788 CCTTGATTGATGTTGTGCTGTAAGAAATTAAGATATCTCATATCTATGATTTGA 3847  
 359 CCTTGATTGATGTTGTGCTGTAAGAAATTAAGATATCTCATATCTATGATTTGA 418

3848 ATGTGTGTAAGAGCTTTGGGCGCGGTAGAGTGGCATGGCTCTTATGAAAGTTAAC 3907  
 419 ATGTGTGTAAGAGCTTTGGGCGCGGTAGAGTGGCATGGCTCTTATGAAAGTTAAC 478

3908 TATTAAGTGGCTTTATGTGCTTGAAGCAATTTCTGAGCGAGACAGTGAAGAG 3967  
 479 TATTAAGTGGCTTTATGTGCTTGAAGCAATTTCTGAGCGAGACAGTGAAGAG 538

3968 TGAATATGATCATGGAAGAACTCAACTATTTAGATTCTGTAATGAAACCCAGTTT 4027  
 539 TGAATATGATCATGGAAGAACTCAACTATTTAGATTCTGTAATGAAACCCAGTTT 598

4028 GTGTTAATATCTGCTGTGAGAACTTAAAGTTCAATATCCCAAGATGCTCAGTGT 4087  
 599 GTGTTAATATCTGCTGTGAGAACTTAAAGTTCAATATCCCAAGATGCTCAGTGT 658

4088 CCATAGTGATTAATTAAGAGCAAGAGAGCAGCGGTGAGAAATTACAATCTGAAGTGA 4147  
 659 CCATAGTGATTAATTAAGAGCAAGAGAGCAGCGGTGAGAAATTACAATCTGAAGTGA 718

4148 AGCTGCTCTCTGTGACCTTTGAGAGTGTCAAGGCTCCGCTCTTGTGAGAGTGA 4207  
 719 AGCTGCTCTCTGTGACCTTTGAGAGTGTCAAGGCTCCGCTCTTGTGAGAGTGA 778

4208 CTTCAGGCTCCCATCATCACTCTTGAATTTTATTTTGTGTTCAAGCTTCTGACT 4267  
 779 CTTCAGGCTCCCATCATCACTCTTGAATTTTATTTTGTGTTCAAGCTTCTGACT 838

4268 TTATGGAACCTTGGCTGTGA 4287  
 |||||

Db 839 TTATGGAACCTTGGCTGTGA 858

RESULT 15  
 AL590428 163577 bp DNA linear PRI 31-JUL-2001  
 LOCUS Human DNA sequence from clone RP11-553A21 on chromosome 6, complete  
 DEFINITION  
 accession AL590428 AC026605  
 version AL590428.7 GI:15072593  
 keywords HTG.  
 SOURCE  
 ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 163577)

REFERENCE  
 AUTHORS Chapman, J.  
 TITLE Direct Submission  
 JOURNAL Submitted (31-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,  
 CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk  
 requests: clonerequests@sanger.ac.uk  
 On Aug 1, 2001 this sequence version replaced gi:15021177.

COMMENT  
 During sequence assembly data is compared from overlapping clones.  
 Where differences are found these are annotated as variations  
 together with a note of the overlapping clone name. Note that the  
 variation annotation may not be found in the sequence submission  
 corresponding to the overlapping clone, as we submit sequences with  
 only a small overlap as described above.  
 This sequence was finished as follows unless otherwise noted: all  
 regions were either double-stranded or sequenced with an alternate  
 chemistry or covered by high quality data (i.e., phred-quality >= 30); an attempt was made to resolve all sequencing problems, such  
 as compressions and repeats; all regions were covered by at least  
 one plasmid subclone or more than one M13 subclone; and the  
 assembly was confirmed by restriction digest. The following  
 abbreviations are used to associate primary accession numbers given  
 in the feature table with their source databases: Em., EMBL; Sw.,  
 SWISSPROT; Tr., TrEMBL; Wp., WORMPEP; Information on the WORMPEP  
 database can be found at  
 http://www.sanger.ac.uk/Projects/C\_elegans/wormpep This sequence  
 was generated from part of bacterial clone contigs of human  
 chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping  
 Group. Further information can be found at  
 http://www.sanger.ac.uk/HGP/Chr6  
 RP11-553A21 is from the library RP11-11.2 constructed by the group  
 of Pieter de Jong. For further details see  
 http://www.chori.org/bacpac/home.htm  
 VECTOR: pBAC3.6  
 IMPORTANT: This sequence is not the entire insert of clone  
 RP11-553A21. It may be shorter because we sequence overlapping  
 sections only once, except for a 100 base overlap.  
 The true right end of clone RP11-553A21 is at 163577 in this  
 sequence. The true left end of clone RP11-525G3 is at 88067 in this  
 sequence. The true right end of clone RP3-197H23 is at 2000 in this  
 sequence.

FEATURES  
 source  
 location/Qualifiers  
 1.163577  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="6"  
 /clone="RP11-553A21"  
 /clone\_11b="RP11-11.2"  
 285..433  
 /note="Alu repeat: matches 1..119 of consensus"  
 434..743  
 /note="Alu repeat: matches 1..311 of consensus"  
 744..919  
 /note="Alu repeat: matches 119..305 of consensus"  
 988..1075  
 /note="L2 repeat: matches 2626..2708 of consensus"  
 1166..1384  
 /note="L1MC5 repeat: matches 7317..7536 of consensus"

repeat\_region 1385. .1690 /note="Alusx repeat: matches 1. .306 of consensus"  
repeat\_region 1691. .1759 /note="L1MCS repeat: matches 7536. .7604 of consensus"  
repeat\_region 1760. .2024 /note="Alu repeat: matches 35. .287 of consensus"  
repeat\_region 2025. .2248 /note="L1MCS repeat: matches 7604. .7887 of consensus"  
repeat\_region 2588. .2706 /note="L1MCS repeat: matches 5389. .5514 of consensus"  
misc\_feature 3686. .4489 /note="CPG island"  
/evidence=not\_experimental  
repeat\_region 3695. .3736 /note="21 copies 2 mer cc 78% conserved"  
repeat\_region 4691. .4996 /note="Alu repeat: matches 3. .306 of consensus"  
repeat\_region 5628. .5977 /note="L1MCS repeat: matches 1. .403 of consensus"  
repeat\_region 5978. .6030 /note="MER5A repeat: matches 44. .97 of consensus"  
repeat\_region 6118. .6474 /note="L1MCS repeat: matches 7398. .7739 of consensus"  
repeat\_region 6475. .6771 /note="Alu repeat: matches 1. .290 of consensus"  
repeat\_region 6772. .7373 /note="L1MCS repeat: matches 6740. .7398 of consensus"  
repeat\_region 7491. .7764 /note="Alu repeat: matches 1. .286 of consensus"  
repeat\_region 7887. .8168 /note="Alusx repeat: matches 1. .287 of consensus"  
repeat\_region 8365. .8430 /note="33 copies 2 mer tc 89% conserved"  
repeat\_region 8431. .8729 /note="Alusx repeat: matches 1. .302 of consensus"  
repeat\_region 8779. .8975 /note="MER3 repeat: matches 129. .323 of consensus"  
repeat\_region 8976. .9292 /note="Alusx repeat: matches 1. .311 of consensus"  
repeat\_region 9293. .9415 /note="MER3 repeat: matches 1. .129 of consensus"  
repeat\_region 9516. .9647 /note="MIR repeat: matches 106. .226 of consensus"  
repeat\_region 9698. .9994 /note="Alusx repeat: matches 1. .295 of consensus"  
repeat\_region 10005. .10048 /note="22 copies 2 mer tt 79% conserved"  
repeat\_region 10050. .10999 /note="SVA repeat: matches 3. .954 of consensus"  
misc\_feature 10622. .11798 /note="CPG island"  
/evidence=not\_experimental  
repeat\_region 11056. .11951 /note="SVA repeat: matches 519. .1386 of consensus"  
repeat\_region 12038. .12152 /note="L2 repeat: matches 2588. .2707 of consensus"  
repeat\_region 12561. .12844 /note="Alusx repeat: matches 3. .287 of consensus"  
repeat\_region 12928. .13220 /note="Alu repeat: matches 2. .295 of consensus"  
repeat\_region 13294. .13417 /note="FLAM C repeat: matches 1. .123 of consensus"  
repeat\_region 13451. .13636 /note="Alusx repeat: matches 1. .189 of consensus"  
repeat\_region 13637. .13697 /note="Alu repeat: matches 1. .63 of consensus"  
repeat\_region 13797. .13925 /note="MER3 repeat: matches 15. .148 of consensus"  
repeat\_region 13926. .14215 /note="Alu repeat: matches 1. .288 of consensus"  
repeat\_region 14216. .14662 /note="MER3 repeat: matches 148. .543 of consensus"  
repeat\_region 14688. .14735

/note="LTR29 repeat: matches 571. .619 of consensus"  
repeat\_region 14832. .15863 /note="L1M4 repeat: matches 5262. .6295 of consensus"  
repeat\_region 15978. .16180 /note="L2 repeat: matches 2208. .2414 of consensus"  
repeat\_region 16387. .16511 /note="L2 repeat: matches 2636. .2734 of consensus"  
repeat\_region 17252. .17283 /note="16 copies 2 mer ac 96% conserved"  
repeat\_region 17372. .17674 /note="Alusx repeat: matches 1. .303 of consensus"  
repeat\_region 18061. .19334 /note="L1M10 repeat: matches 5026. .6320 of consensus"  
repeat\_region 19352. .19437 /note="L1M3C repeat: matches 734. .820 of consensus"  
repeat\_region 19545. .19620 /note="L1M3C repeat: matches 922. .997 of consensus"  
repeat\_region 20319. .20442 /note="Alu repeat: matches 179. .305 of consensus"  
repeat\_region 20921. .21277 /note="L1M10 repeat: matches 1. .364 of consensus"  
repeat\_region 21480. .21775 /note="Alusx repeat: matches 1. .296 of consensus"  
repeat\_region 21799. .21916 /note="FLAM C repeat: matches 1. .118 of consensus"  
repeat\_region 22546. .22713 /note="Alusx repeat: matches 87. .262 of consensus"  
repeat\_region 22895. .23001 /note="L1M2 repeat: matches 5827. .5924 of consensus"  
repeat\_region 23005. .23303 /note="Alusx repeat: matches 14. .312 of consensus"  
repeat\_region 23429. .23525 /note="Alusx repeat: matches 213. .309 of consensus"  
repeat\_region 23528. .23628 /note="Alusx repeat: matches 12. .112 of consensus"  
repeat\_region 23626. .23695 /note="Alusx repeat: matches 62. .142 of consensus"  
repeat\_region 23696. .23996 /note="Alusx repeat: matches 1. .304 of consensus"  
repeat\_region 23997. .24222 /note="Alusx repeat: matches 142. .309 of consensus"  
repeat\_region 24243. .24416 /note="Alusx repeat: matches 127. .300 of consensus"  
repeat\_region 24427. .24734 /note="Alusx repeat: matches 1. .300 of consensus"  
repeat\_region 24807. .25158 /note="MIR1 repeat: matches 1. .363 of consensus"  
repeat\_region 25179. .25484 /note="Alusx repeat: matches 1. .299 of consensus"  
repeat\_region 25494. .25519 /note="13 copies 2 mer at 100% conserved"  
repeat\_region 25614. .25730 /note="FLAM C repeat: matches 2. .118 of consensus"  
repeat\_region 26958. .27263 /note="Alu repeat: matches 7. .311 of consensus"  
repeat\_region 27497. .27814 /note="L1M4 repeat: matches 4499. .4833 of consensus"  
repeat\_region 27866. .27943 /note="L1M10 repeat: matches 410. .492 of consensus"  
  
Query Match 5.4%; Score 232; DB 9; Length 163577;  
Best Local Similarity 100.0%; Pred. No. 6.9e-40;  
Matches 223; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 277 CTACCTGTGAACAGTGCAGATGAGATTATGAGCTACGCTTAACCGAGTACCCAGAT 336  
DB 80550 CTACCTGTGAACAGTGCAGATGAGATTATGAGCTACGCTTAACCGAGTACCCAGAT 80609  
QY 337 GAGATTTTATTCCTATATGATACCGCTTATATGAGCCAGAGATATCTGCTTC 396  
DB 80610 GAGATTTTATTCCTATATGATACCGCTTATATGAGCCAGAGATATCTGCTTC 80669  
QY 397 ATTCAACAGACAGCGCTTATACAGCCAAAGAGAGTGAAGTTTGCATTTTACA 456

Db 80670 ATTCAACAGACAGGCGCTTATACAGCCAAAGCAAGAGTGAAGTTCCGATTGTTACA 80729  
Qy 457 CTCTTCTCAGATTTTAAGCCTTACAAACCTTTTAAACATTCTCATTAAG 508  
Db 80730 CTCTTCTCAGATTTTAAGCCTTACAAACCTTTTAAACATTCTCATTAAG 80781

Search completed: January 16, 2004, 18:48:41  
Job time : 15013 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd

OM nucleic - nucleic search, using sw model

Run on: January 16, 2004, 12:36:15 ; Search time 1017 Seconds  
(without alignments)  
11379.047 Million cell updates/sec

Title: US-10-020-095-3

Sequence: 1 atgcagggccaccgctcct.....ttatggaactttgctgtga 4287

Scoring table: IDENTITY\_NUC

Gapor 10.0 , Gapext 1.0

Searched: 2552756 beqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

Post-processing: Minimum Match 0%

## Listing first 45 summaries

Database : N\_Geneseq\_19Jun03:\*

1:	/SDSI/gcgdatr/genseeq/genseqnm-emb1/NA1980.DAT.*
2:	/SDSI/gcgdatr/genseeq/genseqnm-emb1/NA1981.DAT.*
3:	/SDSI/gcgdatr/genseeq/genseqnm-emb1/NA1982.DAT.*
4:	/SDSI/gcgdatr/genseeq/genseqnm-emb1/NA1983.DAT.*
5:	/SDSI/gcgdatr/genseeq/genseqnm-emb1/NA1984.DAT.*
6:	/SDSI/gcgdatr/genseeq/genseqnm-emb1/NA1985.DAT.*
7:	/SDSI/gcgdatr/genseeq/genseqnm-emb1/NA1986.DAT.*
8:	/SDSI/gcgdatr/genseeq/genseqnm-emb1/NA1987.DAT.*
9:	/SDSI/gcgdatr/genseeq/genseqnm-emb1/NA1988.DAT.*
10:	/SDSI/gcgdatr/genseeq/genseqnm-emb1/NA1989.DAT.*
11:	/SDSI/gcgdatr/genseeq/genseqnm-emb1/NA1990.DAT.*
12:	/SDSI/gcgdatr/genseeq/genseqnm-emb1/NA1991.DAT.*
13:	/SDSI/gcgdatr/genseeq/genseqnm-emb1/NA1993.DAT.*
14:	/SDSI/gcgdatr/genseeq/genseqnm-emb1/NA1995.DAT.*
15:	/SDSI/gcgdatr/genseeq/genseqnm-emb1/NA1994.DAT.*
16:	/SDSI/gcgdatr/genseeq/genseqnm-emb1/NA1995.DAT.*
17:	/SDSI/gcgdatr/genseeq/genseqnm-emb1/NA1996.DAT.*
18:	/SDSI/gcgdatr/genseeq/genseqnm-emb1/NA1997.DAT.*
19:	/SDSI/gcgdatr/genseeq/genseqnm-emb1/NA1998.DAT.*
20:	/SDSI/gcgdatr/genseeq/genseqnm-emb1/NA1999.DAT.*
21:	/SDSI/gcgdatr/genseeq/genseqnm-emb1/NA2000.DAT.*
22:	/SDSI/gcgdatr/genseeq/genseqnm-emb1/NA2001.DAT.*
23:	/SDSI/gcgdatr/genseeq/genseqnm-emb1/NA2001B.DAT.*
24:	/SDSI/gcgdatr/genseeq/genseqnm-emb1/NA2002.DAT.*
25:	/SDSI/gcgdatr/genseeq/genseqnm-emb1/NA2003.DAT.*

**Pred. No.** is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	4281.2	99.9	4369	25	AAD49434	Human r150 gene #1
2	4224.4	98.5	4761	24	ABD79964	Human C109 K1 pro
3	4224.4	98.5	5882	25	AAD49440	Human blood cell s
4	4224.4	98.5	5895	24	ABD79966	Human C109 K1-H7
5	4223.4	98.5	5883	25	AAD49435	Human r150 DNA #1.
6	4222.8	98.5	4761	24	ABD79965	Human C109 K1 var
7	4222.8	98.5	5895	24	ABD79967	Human C109 K1-H7
8	4221.4	98.5	4335	24	AAL49615	Human platelet all

9	4219.8	98.4	4435	24	AAAE2916	Human platelet all
10	4219.6	98.4	4473	21	AAAE6019	Hydrophobic domain
11	4215.6	98.4	4335	21	AAAE6019	Hydrophobic domain
12	4140.2	96.6	4146	25	AAAD9436	Human blood cell s
13	4082.4	95.2	4197	25	AAAD9437	Human r150 DNA #2.
14	3187.4	74.4	3535	24	ABQ79968	Human CD109 K15 pr
15	3185.8	74.3	3535	24	ABQ79969	Human CD109 K15 va
16	1017	23.7	1448	24	ABIE90478	Human polynucleoti
17	975.4	22.8	1459	22	ABAO8885	Human secreted prote
18	812.2	18.9	1403	24	ABAV77330	Alpha-1 proteinase
19	730.8	17.0	1300	22	AAEF17755	Human protease and
20	541.6	12.6	875	20	AAAX0449	Human secreted pro
21	525	12.2	821	25	ABXO8883	Angiogenesis-absoo
22	497.6	11.6	794	20	AAZ15521	Human gene express
23	331.6	7.7	690	23	AAAF0864	Human gene express
24	300	7.0	300	20	AAAZ13539	Human gene express
25	272	6.3	354	25	ABXA47702	Bovine EST associat
26	175	4.1	744	20	AAAG16725	Human gene express
27	174	4.1	4527	22	AAAG0951	C. elegans alpha-2
28	174	4.1	4560	22	AAAG0952	C. elegans alpha-2
29	172.6	4.0	2280	23	ABEL20611	Drosophila melanog
30	171.6	4.0	2485	23	ABEL12153	Drosophila melanog
31	152.2	3.6	3612	23	ABLI19705	Drosophila melanog
32	143.8	3.4	6324	23	ABLI19704	Drosophila melanog
33	127.2	3.0	5611	23	ABEL20610	Drosophila melanog
34	127.2	3.0	6979	23	ABLI12152	Drosophila melanog
35	126.8	3.0	4615	23	AAAG92193	DNA encoding novel
36	126.8	3.0	4615	23	ABAN97232	Gene #3730 used to
37	126.8	3.0	4629	24	AAAG92192	DNA encoding novel
38	126.8	3.0	5816	21	AACT44484	Human ORFX ORF39 p
39	125.2	2.9	4487	24	ABEK2035	DNA encoding novel
40	125.2	2.9	4488	24	ABEK59337	Human alpha-2-macr
41	120.6	2.8	4660	24	ABQ93897	Human pregnancy zc
42	120.6	2.8	2608	24	AALE9817	Human platelet all
43	119.8	2.8	4501	25	ABXY0464	DNA encoding human
44	119.4	2.8	1140	22	ABAO6656	Human CDNA SEQ ID
45	119.4	2.8	1140	22	ABAO6859	Human polynucleoti

## ALIGNMENTS

RESULT 1  
AAD49434  
ID--AAD49434 standard; DNA; 4369 BP

AC AAD49434;

DT 24-MAR-2003 (first entry)

Human r150 gene #1.

KW Human; GPI-anchored TGF-beta1 binding protein; r150 protein; cancer;

KW therapy; gene; ds

OS Homo sapiens

FH	Key	Location/Qualifiers
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

ET

```

/|transl_except= (pos:2132..2134, aa:Xaa)

```

[illegible]

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

2 XX  
2  
4  
g  
g  
o  
o  
o  
o  
o  
o  
o  
o  
o  
o  
o  
o

2000

PR 14-FEB-2002; 2002US-356163P.

XX (UTMC-) UNIV MCGILL.  
 PA Philip A, Tam B;  
 PI WPI: 2003-093100/08.  
 XX P-PSDB; AAE32012.  
 DR Novel transforming growth factor (TGF)-beta 1 binding reagent which  
 PT comprises r150 protein which acts as accessory receptor of TGF-beta,  
 PT useful for negatively modulating TGF-beta activity, and thus for  
 PT treating cancer -  
 XX Claim 7; Page 91-93; 127pp; English.  
 PS The invention relates to novel transforming growth factor (TGF)-beta1  
 CC binding reagent which comprises a glycosylphosphatidylinositol (GPI)-  
 CC anchored TGF-beta1 binding protein referred to as r150 which acts as  
 CC an accessory receptor of TGF-beta. The invention is used for negatively  
 CC modulating TGF-beta activity, and thus for treating conditions  
 CC characterised by overproduction of TGF-beta, such as cancer. Antisense  
 CC molecules of the invention are used for increasing TGF-beta availability  
 CC and increase graft success. The present sequence is a gene encoding  
 CC human r150 protein.  
 CC  
 SQ Sequence 4369 BP; 1301 A; 842 C; 932 G; 1293 T; 1 other;

Query Match 99.9%; Score 4281.2; DB 25; Length 4369;  
 Basic Local Similarity 99.9%; Pred. No. 0;  
 Matches 4283; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGCAAGGCGCCACGCTCTGACCGCCGACCTCTCTGCTGTCGACCGCGCGCTG 60  
 DB 26 ATGCAAGGCGCCACGCTCTGACCGCGCCGACCTCTCTGCTGTCGACCGCGCGCTG 85  
 QY 61 GCCGTGCTCCCGGCGCTCGGTTTCTGTGTGACAGCGCCGAGATCATGAGCCCGAGA 120  
 DB 86 GCCGTGCTCCCGGCGCTCGGTTTCTGTGTGACAGCGCCGAGATCATGAGCCCGAGA 145  
 QY 121 AATGTGATATTTGGGTGAGACTTCTGACACATGCGCTTTCACAGGTGACTGTGAAGCG 180  
 DB 146 AATGTGATATTTGGGTGAGACTTCTGACACATGCGCTTTCACAGGTGACTGTGAAGCG 205  
 QY 181 GAGCTGCTCAAGACAGCATCAACTCACTCTCTCTGCTGTAAGCAAGAGAGTCTT 240  
 DB 206 GAGCTGCTCAAGACAGCATCAACTCACTCTCTCTGCTGTAAGCAAGAGAGTCTT 265  
 QY 241 GAAAAAGGCTCTTTTAAGACACTTACTTTCATCACTACTGTAACAGTGCAGATGAG 300  
 DB 266 GAAAAAGGCTCTTTTAAGACACTTACTTTCATCACTACTGTAACAGTGCAGATGAG 325  
 QY 301 ATTATGAGCTACGTGTACCGGACGTACCGAGATGAGATTTATTTCTTAATAGTACC 360  
 DB 326 ATTATGAGCTACGTGTACCGGACGTACCGAGATGAGATTTATTTCTTAATAGTACC 385  
 QY 361 CGCTTATCATTTGAGACCAAGAAATATCTGCTTCACTTAAGACAGCAAGGCTTTATC 420  
 DB 386 CGCTTATCATTTGAGACCAAGAAATATCTGCTTCACTTAAGACAGCAAGGCTTTATC 445  
 QY 421 AAGCCAAAGCAAGAGTGAAGTTTCGACTTGTACACTCTTCTGATTTTAAGCCTTAC 480  
 DB 446 AAGCCAAAGCAAGAGTGAAGTTTCGACTTGTACACTCTTCTGATTTTAAGCCTTAC 505  
 QY 481 AAAAAGCTCTTTAAACATTTCTATTAAAGACCCCAATCAATTTGATTCACAGTGTG 540  
 DB 506 AAAAAGCTCTTTAAACATTTCTATTAAAGACCCCAATCAATTTGATTCACAGTGTG 565  
 QY 541 TCACAAACAAAGTGAATTTGAGATTTTCAAACTTTTGAATCTTCCATCCATA 600  
 DB 566 TCACAAACAAAGTGAATTTGAGATTTTCAAACTTTTGAATCTTCCATCCATA 625  
 QY 601 CTGTGTGATGTGTATTTCAAGTTCAAGTGAATGACAGACATTTATCATCATTTTCA 660

DB 626 CTGTGTGATGTGTATTTCAAGTTCAAGTGAATGACAGACATCATCATTTTCA 685  
 QY 661 GTTTCAGATATGATTTACCAAAATTTGAAGTACTTTTGACAGACCATATATTTGCT 720  
 DB 686 GTTTCAGATATGATTTACCAAAATTTGAAGTACTTTTGACAGACCATATATTTGCT 745  
 QY 721 ATGAATTTGAAGATTTAAATGGTACATCAAGCAAGATATCATATGGGAAGCAGTG 780  
 DB 746 ATGAATTTGAAGATTTAAATGGTACATCAAGCAAGATATCATATGGGAAGCAGTG 805  
 QY 781 AAGGAGAGCTACCGTATCATTTTACCTTTATCTTTTGGGAAAGAAATATTT 840  
 DB 806 AAGGAGAGCTACCGTATCATTTTACCTTTATCTTTTGGGAAAGAAATATTT 865  
 QY 841 AAAAAACATTTAAGATTAATGATCTGCAAACTTCTTTTATGATGAGAGATGAAA 900  
 DB 866 AAAAAACATTTAAGATTAATGATCTGCAAACTTCTTTTATGATGAGAGATGAAA 925  
 QY 901 AATGTATGATTTCTTCAATGAGACTTTTGAATACCTGATCTATCTTCCCTGACCA 960  
 DB 926 AATGTATGATTTCTTCAATGAGACTTTTGAATACCTGATCTATCTTCCCTGACCA 985  
 QY 961 GTTGAATTTTAAACCAAGTGAACAGATCATGTTACAGTTATTTCAAGAAATGATACCT 1020  
 DB 986 GTTGAATTTTAAACCAAGTGAACAGATCATGTTACAGTTATTTCAAGAAATGATACCT 1045  
 QY 1021 AATGTGTTCTTCAAGACATGATTTATCATGATTTGATTTTGAATATATGATCTG 1080  
 DB 1046 AATGTGTTCTTCAAGACATGATTTATCATGATTTGATTTTGAATATATGATCTG 1105  
 QY 1081 AAGCCATCTCTCAACTTCAAGCAAGCACTGTGAAGTAACTGTGCTGATGCAACCACTG 1140  
 DB 1106 AAGCCATCTCTCAACTTCAAGCAAGCACTGTGAAGTAACTGTGCTGATGCAACCACTG 1165  
 QY 1141 ACTCTTGAAGAAAGAAATATATGATCTTAACTGATCAACAGAGAAATCTATCTGAG 1200  
 DB 1166 ACTCTTGAAGAAAGAAATATATGATCTTAACTGATCAACAGAGAAATCTATCTGAG 1225  
 QY 1201 TACTGAGGAGGATCTTAAAGTGAATGAGAAATGGAAGCTGTGTGAGAAATTAAT 1260  
 DB 1226 TACTGAGGAGGATCTTAAAGTGAATGAGAAATGGAAGCTGTGTGAGAAATTAAT 1285  
 QY 1261 ACTGTCCCCCAAGTGAACCTTTTAAGATTGAATTCACATCTGTGAGATTCAGTGAG 1320  
 DB 1286 ACTGTCCCCCAAGTGAACCTTTTAAGATTGAATTCACATCTGTGAGATTCAGTGAG 1345  
 QY 1321 CTACAGTTGAAGGCTTATTTCTTGTGATGAAAGTACATGCACTGATCTGTTT 1380  
 DB 1346 CTACAGTTGAAGGCTTATTTCTTGTGATGAAAGTACATGCACTGATCTGTTT 1405  
 QY 1381 AAGTCTCTAGTGAACATTAATCCAACTTAAACCAAGATGAGAAATTAAGTGGGA 1440  
 DB 1406 AAGTCTCTAGTGAACATTAATCCAACTTAAACCAAGATGAGAAATTAAGTGGGA 1465  
 QY 1441 TCGCCTTTTGAAGTGTGTTGTTAGTGGCAAAACGATTTGAAGAGTTAAGTATAGTGA 1500  
 DB 1466 TCGCCTTTTGAAGTGTGTTGTTAGTGGCAAAACGATTTGAAGAGTTAAGTATAGTGA 1525  
 QY 1501 GTATCCAGGGGACAGTTGTGCTGTAGAGAAACAAATTTCAACAAATGTTCTTTTACA 1560  
 DB 1526 GTATCCAGGGGACAGTTGTGCTGTAGAGAAACAAATTTCAACAAATGTTCTTTTACA 1585  
 QY 1561 CCAAGAAATTTCTTGAACCTCCAAAGGCTGTGATTTGATTTAATTTGAAGTGAATGG 1620  
 DB 1586 CCAAGAAATTTCTTGAACCTCCAAAGGCTGTGATTTGATTTAATTTGAAGTGAATGG 1645  
 QY 1621 GAAATTAAGTGAATTTCTTAAATTTCTGTTTCACTGTTTAAATTAAGTAAAG 1680  
 DB 1646 GAAATTAAGTGAATTTCTTAAATTTCTGTTTCACTGTTTAAATTAAGTAAAG 1705  
 QY 1681 GTATATTTGAATTAAGTGAAGCTGAAACATCTGAGAAAGTCTCTTGAAGATCTGTG 1740  
 DB 1706 GTATATTTGAATTAAGTGAAGCTGAAACATCTGAGAAAGTCTCTTGAAGATCTGTG 1765

QY 1741 AACAGCCTGACTCCATAGTGGGATTTGAGCTGTGACAAAAGTGTGAATCTGAAT 1800  
DB 1766 AACAGCCTGACTCCATAGTGGGATTTGAGCTGTGACAAAAGTGTGAATCTGAAT 1825  
QY 1801 GCTCTAATGATTTTCAATGGAATAATGTGTCCATGAGTGGAACTTTATAACAGAGA 1860  
DB 1826 GCTCTAATGATTTTCAATGGAATAATGTGTCCATGAGTGGAACTTTATAACAGAGA 1885  
QY 1861 TATTATTAGGCAATGTCATGTAATCTTTGAGCTTTCAAGAAATGTGACCTGGGTA 1920  
DB 1886 TATTATTAGGCAATGTCATGTAATCTTTGAGCTTTCAAGAAATGTGACCTGGGTA 1945  
QY 1921 TTGACAGATGCAAACTTCAGAGATTAATTTGATGTGTGTTATGACAAATGAGAAAT 1980  
DB 1946 TTGACAGATGCAAACTTCAGAGATTAATTTGATGTGTGTTATGACAAATGAGAAAT 2005  
QY 1981 GCTGAGAGGTTTATGAGAGAAATGAGACATATTGTGATATTCATGACTTTCTTTG 2040  
DB 2006 GCTGAGAGGTTTATGAGAGAAATGAGACATATTGTGATATTCATGACTTTCTTTG 2065  
QY 2041 GGTAGACATCCATGTCGGAAGATTTTCCAGAGACTTGATTTGCTAGACACCAAC 2100  
DB 2066 GGTAGACATCCATGTCGGAAGATTTTCCAGAGACTTGATTTGCTAGACACCAAC 2125  
QY 2101 ATGGGTTACAGGATTTTCCAGAAATTTGAAGTAATCTGATCTGATTTCTATCACTTTGG 2160  
DB 2126 ATGGGTTACAGGATTTTCCAGAAATTTGAAGTAATCTGATCTGATTTCTATCACTTTGG 2185  
QY 2161 GTGGCTACTGTGTTTGTGATCTCTGAGACCTGGGCTTTGACATCACTCTGAGT 2220  
DB 2186 GTGGCTACTGTGTTTGTGATCTCTGAGACCTGGGCTTTGACATCACTCTGAGT 2245  
QY 2221 GAGCTCCAGCCTTCCAACTTTTCAATTTTGTGAACTTCCCTGACTGTGATACGA 2280  
DB 2246 GAGCTCCAGCCTTCCAACTTTTCAATTTTGTGAACTTCCCTGACTGTGATACGA 2305  
QY 2281 GGTGAAGATTTGCTTTGAAATTAATCTATTTCAATTTTGAAGATGCCACTGAGTT 2340  
DB 2306 GGTGAAGATTTGCTTTGAAATTAATCTATTTCAATTTTGAAGATGCCACTGAGTT 2365  
QY 2341 AAGGTATCATTTAGAAAAGTGAACAATTTGATTTCTAATGACTTCAAGTGAATTAAT 2400  
DB 2366 AAGGTATCATTTAGAAAAGTGAACAATTTGATTTCTAATGACTTCAAGTGAATTAAT 2425  
QY 2401 GGCACAGGCGACAGAGACCTTCTGTGTTCCAGAGAGATGGGCAACTGTTCTTTT 2460  
DB 2426 GGCACAGGCGACAGAGACCTTCTGTGTTCCAGAGAGATGGGCAACTGTTCTTTT 2485  
QY 2461 CCCATCAGGCGAACACATCTGGGAGAAATTCCTATCAGATCAGCTCTTTCAACCACT 2520  
DB 2486 CCCATCAGGCGAACACATCTGGGAGAAATTCCTATCAGATCAGCTCTTTCAACCACT 2545  
QY 2521 GCTTCTGATGTGTCAACCAAGATTTTGTAAAGGCTGAAGAAATAGAAAATCATAT 2580  
DB 2546 GCTTCTGATGTGTCAACCAAGATTTTGTAAAGGCTGAAGAAATAGAAAATCATAT 2605  
QY 2581 TCACATTCATCTTATTAAGTGTGACATAGCTACAGAGTACCTGAAAATTTTG 2640  
DB 2606 TCACATTCATCTTATTAAGTGTGACATAGCTACAGAGTACCTGAAAATTTTG 2665  
QY 2641 AGTTTCTATTCCTCTAATACAGTACGTCGAGTGAAGAAAGTTGAGATCTGCAAT 2700  
DB 2666 AGTTTCTATTCCTCTAATACAGTACGTCGAGTGAAGAAAGTTGAGATCTGCAAT 2725  
QY 2701 GAGAGATTTCTTGCTCTTCCATCAATGCTTAGCTCATTTGATGCTTATGCTG 2760  
DB 2726 GAGAGATTTCTTGCTCTTCCATCAATGCTTAGCTCATTTGATGCTTATGCTG 2785  
QY 2761 TGTGTGGAACAGAACTGATAAATTTTGTCTCAAAATATTACATTTTGTGATTAATCTGACT 2820  
DB 2786 TGTGTGGAACAGAACTGATAAATTTTGTCTCAAAATATTACATTTTGTGATTAATCTGACT 2845

QY 2821 AAAAAAGAACATGACAGATTAATTTGAAAGAAAAAGCTTTTCATTTATGAGGCAAGT 2880  
DB 2846 AAAAAAGAACATGACAGATTAATTTGAAAGAAAAAGCTTTTCATTTATGAGGCAAGT 2905  
QY 2881 TACAGAGAAAGATTTCTTATCAGAGGGAAGATGGCTTTTCAAGTCTTTTGGAAATTAAT 2940  
DB 2906 TACAGAGAAAGATTTCTTATCAGAGGGAAGATGGCTTTTCAAGTCTTTTGGAAATTAAT 2965  
QY 2941 GACCCCTTGGGAGACATTTGTTGTCAGCTTTTGTGTTTGAAGTGTTCCTTGAAGCCGAT 3000  
DB 2966 GACCCCTTGGGAGACATTTGTTGTCAGCTTTTGTGTTTGAAGTGTTCCTTGAAGCCGAT 3025  
QY 3001 CCTTACATGATTAATGATCAGAAATGTGTTTACAGAAATATACACTTGTGCTTAAAGACAT 3060  
DB 3026 CCTTACATGATTAATGATCAGAAATGTGTTTACAGAAATATATCACTTGTGCTTAAAGACAT 3085  
QY 3061 CAGAAATTCAGACGTTGAATTTTGGATTCAGAGAAAGTGTATCATAGTACCTTCAAGT 3120  
DB 3086 CAGAAATTCAGACGTTGAATTTTGGATTCAGAGAAAGTGTATCATAGTACCTTCAAGT 3145  
QY 3121 GGCATTAAGTCCAGTACACTTACAGCCTATATTTGTAATCTTCTCTGGGATATAGA 3180  
DB 3146 GGCATTAAGTCCAGTACACTTACAGCCTATATTTGTAATCTTCTCTGGGATATAGA 3205  
QY 3181 AAGTATCAGCCTTAACTTGAATGTGCAAGAGTCTATCCATTTTGTGAGTGAATTCAGT 3240  
DB 3206 AAGTATCAGCCTTAACTTGAATGTGCAAGAGTCTATCCATTTTGTGAGTGAATTCAGT 3265  
QY 3241 AGAGGAATTTTCAAGCAATTAATCTAGCCTTATTAATTCATATGATGTCAATGAGTGGG 3300  
DB 3266 AGAGGAATTTTCAAGCAATTAATCTAGCCTTATTAATTCATATGATGTCAATGAGTGGG 3325  
QY 3301 AGTCTTAAGGAGGAAGAGCTTTGATATGTCGACTTGGAGAGAGAAAGAAAGTGGC 3360  
DB 3326 AGTCTTAAGGAGGAAGAGCTTTGATATGTCGACTTGGAGAGAGAAAGAAAGTGGC 3385  
QY 3361 ATGCAATTTGCGGTGTCAATCAGAGTCCAACTTTCTGACTCTGACAGCAGCTCCCTG 3420  
DB 3386 ATGCAATTTGCGGTGTCAATCAGAGTCCAACTTTCTGACTCTGACAGCAGCTCCCTG 3445  
QY 3421 GATATTGAAGTTGACGCTTATGACCTGTCTTCAACTTTTCAATTTTCAACTTTCTGAG 3480  
DB 3446 GATATTGAAGTTGACGCTTATGACCTGTCTTCAACTTTTCAACTTTTCAACTTTCTGAG 3505  
QY 3481 GGAATCCCAATTAAGAGTGTGCTAAGAGGCAAGAAATTAAGTGGGTTTGGATCT 3540  
DB 3506 GGAATCCCAATTAAGAGTGTGCTAAGAGGCAAGAAATTAAGTGGGTTTGGATCT 3565  
QY 3541 ACTCAGATACCACTGTGGCTTTAAAGGCTGTGTGAATTTGACAGCCTTATGATTAACA 3600  
DB 3566 ACTCAGATACCACTGTGGCTTTAAAGGCTGTGTGAATTTGACAGCCTTATGATTAACA 3625  
QY 3601 GAAAGCAAAATATCAAGTGAACGTTGACGGGCTTACCTCAACAAAGTCTTGTGTG 3660  
DB 3626 GAAAGCAAAATATCAAGTGAACGTTGACGGGCTTACCTCAACAAAGTCTTGTGTG 3685  
QY 3661 GTACAGCCAAATGCGAGTAAATTTCCGCAAAATGTTTGTGATTTGCTAATTTGACGTC 3720  
DB 3686 GTACAGCCAAATGCGAGTAAATTTCCGCAAAATGTTTGTGATTTGCTAATTTGACGTC 3745  
QY 3721 AATGTTGATTAATATGTAAGGCTTCTGGGCTTCTTGAAGAACAGAAATCTTATCCAAAT 3780  
DB 3746 AATGTTGATTAATATGTAAGGCTTCTGGGCTTCTTGAAGAACAGAAATCTTATCCAAAT 3805  
QY 3781 CAAAGAGCCTTTGATTTGATGTTGCTGTAAAGAAATTAAGATGATCTCAATCATGTG 3840  
DB 3806 CAAAGAGCCTTTGATTTGATGTTGCTGTAAAGAAATTAAGATGATCTCAATCATGTG 3865  
QY 3841 GATTTGATGTGTGACAGCTTTTCCGCCCCGGGTAGAGATGCGATGCTTATAGAA 3900  
DB 3866 GATTTGATGTGTGACAGCTTTTCCGCCCCGGGTAGAGATGCGATGCTTATAGAA 3925  
QY 3901 GTTAACTTAATAGGCTTTATGTTGCTTCAAGAAATTTCTCTGAGGAGACAGT 3960

```

Db      3926 GTTAACTTAAAGTGGCTTTATGGTCTCAGAGCAATTTCTGAGCGACAGAGTGT
Qy      3961 AAGAAAGTGAATATGATCATGAAAACTCAACCTATTTAGATTCGTAAATGAAC
Db      3966 AAGAAAGTGAATATGATCATGAAAACTCAACCTATTTAGATTCGTAAATGAAC
Qy      4021 CAGTTTGTGTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAATACCAAGATGCT
Db      4046 CAGTTTGTGTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAATACCAAGATGCT
Qy      4081 TCAGTGTCCATAGTGGATTACTATGAGCCAAAGAGACAGCGGTGAGAGATTACACTT
Db      4106 TCAGTGTCCATAGTGGATTACTATGAGCCAAAGAGACAGCGGTGAGAGATTACACTT
Qy      4141 GAAGTGAAGCTGCTCCTGAGACCTTGAGATGTCAGAGGCTGCCCTTGTGAG
Db      4166 GAAGTGAAGCTGCTCCTGAGACCTTGAGATGTCAGAGGCTGCCCTTGTGAG
Qy      4201 GATGAGCTTCAAGCTCCCATCATCATCTTTCAGTCAATTTTATTTCTGTTCAAGCTT
Db      4226 GATGAGCTTCAAGCTCCCATCATCATCTTTCAGTCAATTTTATTTCTGTTCAAGCTT
Qy      4261 CTGTACTTTATGAACTTTGCTGTGA
Db      4286 CTGTACTTTATGAACTTTGCTGTGA

```

## RESULT 2

ABQ79964 ID ABQ79964 standard; cDNA; 4761 BP.

AC ABQ79964;

DT 23-DEC-2002 (first entry)

DE Human CD109 K1 protein encoding cDNA.

KW CD109; alpha2 macroglobulin; thioester; cerebroprotective; cardiant;  
immunopressive; haemostatic; anticoagulant; thrombolytic; human;  
cardiovascular; vasotropic; gene therapy; CD109 K1; gene; ss.

OS Homo sapiens.

Key Location/Qualifiers

FT CDS 113..4450

FT /\*tag= a

FT /product= "CD109 K1"

XX MO200270696-A2.

XX 12-SEP-2002.

XX 07-MAR-2002; 2002WO-CA00292.

XX 07-MAR-2001; 2001US-273814P.

XX (SCHU//) SCHUH A.

XX (SUTH//) SUTHERLAND R D.

XX Schuh A, Sutherland RD;

XX WPI: 2002-713450/77.

XX P-PSDB: ABB82165.

XX New CD109 nucleic acids and polypeptides, useful in gene therapy,

XX particularly for treating strokes, myocardial infarctions, thrombosis,

XX chromocyclopnea, autoimmune diseases, or organ or bone marrow

XX transplantation -

XX Claim 1; Fig 1a; 156pp; English.

XX The invention relates to isolated nucleic acid molecules encoding CD109

CC polypeptides. These nucleic acid molecules include the human cDNA  
CC sequences comprising CD109 K1, CD109 K1-H7, CD109 K1S or their variants.  
CC CD109 is a novel member of the alpha2 macroglobulin (alpha2M/C3, C4, C5  
CC family of thioester-containing proteins. The CD109 polypeptides can be  
CC expressed by standard recombinant methodology. The CD109 nucleic acid,  
CC CD109 polypeptide or its mimetic, or the anti-CD109 antibody is useful  
CC for treating or detecting a disease or disorder, e.g. conditions  
CC associated with endothelial activation, platelet activation, activation  
CC of the coagulation or fibrinolytic systems, activation of T lymphocytes  
CC and of the complement system, quantitative or qualitative abnormalities  
CC of platelet function, increased or impaired platelet aggregation and  
CC activation, increased or impaired activation of the coagulation and/or  
CC fibrinolytic systems, or impaired or increased immune activation. These  
CC are also useful for treating cardiovascular disorders, stroke, myocardial  
CC infarction, thrombosis, embolism, peripheral vascular disease,  
CC thrombocytopenia, thrombocytopenia, autoimmune diseases, organ  
CC transplantation, or bone marrow transplantation. The CD109 nucleic acids  
CC or their antisense nucleotide sequence are useful in gene therapy for  
CC treating these conditions. The present sequence represents the human  
CC CD109 K1 cDNA sequence.

XX Sequence 4761 BP; 1420 A; 912 C; 1016 G; 1413 T; 0 other;

Query Match 98.5%; Score 4224.4; DB 24; Length 4761;  
Best Local Similarity 98.8%; Pred. No. 0;  
Matches 4286; Conservative 0; Mismatches 1; Indels 51; Gaps 1;

```

Qy      1 ATGAGGGGCGCCACGCTCCGACCGCGCCACCTCCTGCTGTCAGCCGCGGCTG
Db      113 ATGAGGGGCGCCACGCTCCGACCGCGCCACCTCCTGCTGTCAGCCGCGGCTG
Qy      61 GCCGTGGCTCCCGGGCTCGGTTCTGTGACAGCCCGAGATCATCAGCCCGAGGA
Db      173 GCCGTGGCTCCCGGGCTCGGTTCTGTGACAGCCCGAGATCATCAGCCCGAGGA
Qy      121 AATGTGACTATTTGGGTGAGCTTCTGGAACACTGCTTACAGGTGACTGGAAGGG
Db      233 AATGTGACTATTTGGGTGAGCTTCTGGAACACTGCTTACAGGTGACTGGAAGGG
Qy      181 GACCTGCTCAAGCAGCATCAAACTCAGTCTGCTGCTGGAACAGAGAGCTTT
Db      293 GACCTGCTCAAGCAGCATCAAACTCAGTCTGCTGCTGGAACAGAGAGCTTT
Qy      241 GAAAAAGGCTCTTTAAGACACTTACTCTTCCATCACTCTCTGAACAGTGCATGAG
Db      353 GAAAAAGGCTCTTTAAGACACTTACTCTTCCATCACTCTCTGAACAGTGCATGAG
Qy      301 ATTATGAGCTACGTGTAACCGGACGTACCGAGATGAGATTTATTTCTTAATAGTACC
Db      413 ATTATGAGCTACGTGTAACCGGACGTACCGAGATGAGATTTATTTCTTAATAGTACC
Qy      361 CGCTTATCATTTGAGACCAAGAGATATGCTTCTTCAATCAACAGCAAGGCTTATAC
Db      473 CGCTTATCATTTGAGACCAAGAGATATGCTTCTTCAATCAACAGCAAGGCTTATAC
Qy      421 AAGCCAAAGCAAGATGAAGTTTCGATTTTACACTCTTCTCAGATTTTAAAGCTTAC
Db      533 AAGCCAAAGCAAGATGAAGTTTCGATTTTACACTCTTCTCAGATTTTAAAGCTTAC
Qy      481 AAAAAGCTTTTAAACATTTCTCATTAAGAACCCCAATCAAAATTTGATCCACAGTGTG
Db      593 AAAAAGCTTTTAAACATTTCTCATTAAGAACCCCAATCAAAATTTGATCCACAGTGTG
Qy      541 TCACAAAGCAAGATGTTGAGAGCTTCCAAACCTTTCAGTATCTTCCATCCCAATA
Db      653 TCACAAAGCAAGATGTTGAGAGCTTCCAAACCTTTCAGTATCTTCCATCCCAATA
Qy      601 CTGTGAGCTGCTATTCATTAAGTGAAGTGAAGACAGACATATATCATCATTTTCAG
Db      713 CTGTGAGCTGCTATTCATTAAGTGAAGTGAAGACAGACATATATCATCATTTTCAG
Qy      661 GTTTCAGAAATGATATACCAAAATTTGAAGTACTTTGACAGACACATTAATTTGTTCT

```



Db 773 GTTACAGATATGATATACCAAAATTTGAAGTACTTTGCAGACACCATATATTTGTTCT 832  
Qy 721 ATGAATCTTAACCAATTAATATGTAACATCAAGGCAAGATATATATGGAAGCCAGTG 780  
Db 833 ATGAATCTTAAGCAATTAATATGTAACATCAAGGCAAGATATATATGGAAGCCAGTG 892  
Qy 781 AAGAGACGTAAGCTTACATTTTACCTTTTACCTTTTGGGGAAGAGAAAATAT 840  
Db 893 AAGAGACGTAAGCTTACATTTTACCTTTTACCTTTTGGGGAAGAGAAAATAT 952  
Qy 841 ACAAACCAATTAAGATTAATGATCTGCAAACTCTCTTTAATGATGAAGATGA 900  
Db 953 ACAAACCAATTAAGATTAATGATCTGCAAACTCTCTTTAATGATGAAGATGA 1012  
Qy 901 AATGATGATGATCTTCAAAATGATCTTCTGAATACCTGATCTATCTTCCCTGACCA 960  
Db 1013 AATGATGATGATCTTCAAAATGATCTTCTGAATACCTGATCTATCTTCCCTGACCA 1072  
Qy 961 GTAGAAATTTTACCAAGTGAAGATCAAGATCAAGTATTTCAAGAAATGTAAGCACT 1020  
Db 1073 GTAGAAATTTTACCAAGTGAAGTGAAGATCAAGATCAAGTATTTCAAGAAATGTAAGCACT 1132  
Qy 1021 AATGATGATGATCTTCAAAATGATCTTCTGAATACCTGATCTATCTTCTGCTTG 1080  
Db 1133 AATGATGATGATCTTCAAAATGATCTTCTGAATACCTGATCTATCTTCTGCTTG 1192  
Qy 1081 AAGCATCTCTCAACTTCAAGCCTCAAGTGAAGTATCTGATGATGGAACCACTG 1140  
Db 1193 AAGCATCTCTCAACTTCAAGCCTCAAGTGAAGTATCTGATGATGGAACCACTG 1252  
Qy 1141 ACTCTTGAAGAAAGAAATATATGATCATTAACATGACACAGAGAACTATATCTGAG 1200  
Db 1253 ACTCTTGAAGAAAGAAATATATGATCATTAACATGACACAGAGAACTATATCTGAG 1312  
Qy 1201 TACTGAGACGATCTTAACATGATGAATCAAGAAATGGAAGCTGTTCAAGAAATATAT 1260  
Db 1313 TACTGAGACGATCTTAACATGATGAATCAAGAAATGGAAGCTGTTCAAGAAATATAT 1372  
Qy 1261 ACTGTCCTCCCAAGTGAAGCTTTTAAGATGAATCCCAATCTGAGAGATCCAGTGAG 1320  
Db 1373 ACTGTCCTCCCAAGTGAAGCTTTTAAGATGAATCCCAATCTGAGAGATCCAGTGAG 1432  
Qy 1321 CTACAGTTGAAGGCTATTTCTTGTAGTAAAGATGATGAGATGAGATGATGATGAT 1380  
Db 1433 CTACAGTTGAAGGCTATTTCTTGTAGTAAAGATGATGAGATGAGATGATGATGAT 1492  
Qy 1381 AAGTCTCTAGTAAACATATCAACTAAACAGAGATGAAGATGAAGATGAAGTGGGA 1440  
Db 1493 AAGTCTCTAGTAAACATATCAACTAAACAGAGATGAAGATGAAGATGAAGTGGGA 1552  
Qy 1441 TCGCTTTTGAAGT 1500  
Db 1553 TCGCTTTTGAAGT 1612  
Qy 1501 GTATCCAGGGGACATTTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1560  
Db 1613 GTATCCAGGGGACATTTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1672  
Qy 1561 CCAGAAATTTCTTGAATCTCAAAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1620  
Db 1673 CCAGAAATTTCTTGAATCTCAAAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1732  
Qy 1621 GAAATTAATGATGATCTTAAATTCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1680  
Db 1733 GAAATTAATGATGATGATCTTAAATTCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1792  
Qy 1681 CTATATTTGAGTAAAGTGAAGCTGAACCATCTGAAGAACTCTCTTGAAGATCTGTG 1740  
Db 1793 CTATATTTGAGTAAAGTGAAGCTGAACCATCTGAAGAACTCTCTTGAAGATCTGTG 1852  
Qy 1741 ACACAGCTGACTCAATGATGAGATTTGATGATGATGATGATGATGATGATGATGAT 1800  
Db 1853 ACACAGCTGACTCAATGATGAGATTTGATGATGATGATGATGATGATGATGATGAT 1912

Qy 1801 GCCTTAATGATATTAACATGAGAAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1860  
Db 1913 GCCTTAATGATATTAACATGAGAAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1972  
Qy 1861 TATTAATTTAGGATGATGATATTTCTTGTGAGTCTTTCAGAAATGTGACTGTGGTA 1920  
Db 1973 TATTAATTTAGGATGATGATATTTCTTGTGAGTCTTTCAGAAATGTGACTGTGGTA 2032  
Qy 1921 TTGACAGATGAGAACTTCAAGAGATTAATGATGATGATGATGATGATGATGATGAT 1980  
Db 2033 TTGACAGATGAGAACTTCAAGAGATTAATGATGATGATGATGATGATGATGATGAT 2092  
Qy 1981 GCTGAGATGATGAGAGAAATGAGAGATTAATGATGATGATGATGATGATGATGATGAT 2040  
Db 2093 GCTGAGATGATGAGAGAAATGAGAGATTAATGATGATGATGATGATGATGATGATGAT 2152  
Qy 2041 GGTAGCACTCAATGTCGAAAGCATTTTCAAGACTTGTGATTTGGCTAGACCAAC 2100  
Db 2153 GGTAGCACTCAATGTCGAAAGCATTTTCAAGACTTGTGATTTGGCTAGACCAAC 2212  
Qy 2101 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2160  
Db 2213 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2272  
Qy 2161 GTGCTACTGATTTTGTGATCTTGAAGACCTGGGTCTTGAATCAACTACTGAGTG 2220  
Db 2273 GTGCTACTGATTTTGTGATCTTGAAGACCTGGGTCTTGAATCAACTACTGAGTG 2332  
Qy 2221 GAGCTCAAGCTTTCACCAATTTTTCATTTTGTGAATCTTCTCTACTCTGTATACGA 2280  
Db 2333 GAGCTCAAGCTTTCACCAATTTTTCATTTTGTGAATCTTCTCTACTCTGTATACGA 2392  
Qy 2281 GGTGAAGATTTGCTTGTGAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2340  
Db 2393 GGTGAAGATTTGCTTGTGAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2452  
Qy 2341 AAGGTAATCTTGAAGAAATGACAAATTTGATATTTCAATGATCTTCAAGTGAATTAAT 2400  
Db 2453 AAGGTAATCTTGAAGAAATGACAAATTTGATATTTCAATGATCTTCAAGTGAATTAAT 2512  
Qy 2401 GCACAGGCAACAGACAGACCTTCTGTGTTCCAGTGAAGATGGGGCAACTGTTCTTTT 2460  
Db 2513 GCACAGGCAACAGACAGACCTTCTGTGTTCCAGTGAAGATGGGGCAACTGTTCTTTT 2572  
Qy 2461 CCATCAGGCAACATCTGGAGAAATTCCTATCACAGTCAAGCTTTTCAACCACT 2520  
Db 2573 CCATCAGGCAACATCTGGAGAAATTCCTATCACAGTCAAGCTTTTCAACCACT 2632  
Qy 2521 GCTTCTGATGCTGTCAACCAATGATGATTTTGAAGGCTGAAGGATGAAGAAATCATAT 2580  
Db 2633 GCTTCTGATGCTGTCAACCAATGATGATTTTGAAGGCTGAAGGATGAAGAAATCATAT 2692  
Qy 2581 TCACATCATCTTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2640  
Db 2693 TCACATCATCTTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2752  
Qy 2641 AGTTCTCATTTCTCTTAATACAGTGACTGGCAGTGAAGAGATTCAGATCACTGCAAT 2700  
Db 2753 AGTTCTCATTTCTCTTAATACAGTGACTGGCAGTGAAGAGATTCAGATCACTGCAAT 2812  
Qy 2701 GGAATGTTCTTGTGCTTCAATCAATGATGATGATGATGATGATGATGATGATGATGAT 2760  
Db 2813 GGAATGTTCTTGTGCTTCAATCAATGATGATGATGATGATGATGATGATGATGATGAT 2872  
Qy 2761 TGTGTGAAGAGACATGATTAATTTGTGCTCAATATTTTCAATTTGATTAATCTGACT 2820  
Db 2873 TGTGTGAAGAGACATGATTAATTTGTGCTCAATATTTTCAATTTGATTAATCTGACT 2932  
Qy 2821 AAAAAAGAAACATGACAGATTAATTTGAAGAAAGCTTTTCAATTTATGAGGCAAGT 2880  
Db 2933 AAAAAAGAAACATGACAGATTAATTTGAAGAAAGCTTTTCAATTTATGAGGCAAGT 2992

OY	2881	TACCAAGAGAACTCTCTATCAGAGGGAAAGATGCGCTTTCAGTGGCTTTGGAAATTAT	2940
Dp	2993	TACAGAGAGAACTTCTATCAGAGGAAAGATGCGCTTTCAGTGGCTTTGGAAATTAT	3052
OY	2941	GACCCCTTCTGGGAGCACTTGGTGTGCAGCTTTTGTTTAAGATGTTTCTTGAAGCCGAT	3000
Dp	3053	GACCCCTTCTGGGAGCACTTGGTGTGCAGCTTTTGTTTAAGATGTTTCTTGAAGCCGAT	3112
OY	3001	CCTTACATAGATATTGATCAGATGTGTTACAGACATACACTTGGCTTAAAGACAT	3060
Dp	3113	CCTTACATAGATATTGATCAGATGTGTTACAGACATACACTTGGCTTAAAGACAT	3172
OY	3061	CAGAAATCCAACGGGAATTTTGGGATCCAGGAAGAGATTCATATGAGACCTCAAGGT	3120
Dp	3173	CAGAAATCCAACGGGAATTTTGGGATCCAGGAAGAGATTCATATGAGACCTCAAGGT	3232
OY	3121	GGCAATAAAGATCCAGTAACTTACACGCTATATTTGAACCTTCTCTCGGATATAGA	3180
Dp	3233	GGCAATAAAGATCCAGTAACTTACACGCTATATTTGAACCTTCTCTCGGATATAGA	3292
OY	3181	AAGTATCAGCCTAACATTTGATGTGCAGAGTCTATCCATTTTTTGGAGTCTGAATTGAGT	3240
Dp	3293	AAGTATCAGCCTAACATTTGATGTGCAGAGTCTATCCATTTTTTGGAGTCTGAATTGAGT	3352
OY	3241	AGAGGAATTTGAGCAACATTAATCTTACCCCTTATTAATTTAGCATTTGTCACTAGTGGG	3300
Dp	3353	AGAGGAATTTGAGCAACATTAATCTTACCCCTTATTAATTTAGCATTTGTCACTAGTGGG	3412
OY	3301	AGCTCTAAGCGAAGAGAGCTTTGAAATATGCTGACTTGGAGAGCGAACAAGAAAGTGGC	3360
Dp	3413	AGCTCTAAGCGAAGAGAGCTTTGAAATATGCTGACTTGGAGAGCGAACAAGAAAGTGGC	3472
OY	3361	ATGCAATTTCTGGGTGTCAATCAGAGTCCAACTTTCTGCACTCTGGCAGCCACGCTCCCTG	3420
Dp	3473	ATGCAATTTCTGGGTGTCAATCAGAGTCCAACTTTCTGCACTCTGGCAGCCACGCTCCCTG	3532
OY	3421	GATATTTGAAGTTGCGAGCCTATGCACTGTCTTCACACTTCTTAACAATTTCAACTTTTGAG	3480
Dp	3533	GATATTTGAAGTTGCGAGCCTATGCACTGTCTTCACACTTCTTAACAATTTCAACTTTTGAG	3592
OY	3481	GGAAATCCCAATTAATGAGGTGGCTGAGGAGCGAAGAAATTAAGCTTGGGTGTTTGCAATCT	3540
Dp	3593	GGAAATCCCAATTAATGAGGTGGCTGAGGAGCGAAGAAATTAAGCTTGGGTGTTTGCAATCT	3652
OY	3541	ACTCAGATATCCACTGTGGCTTTAAAGGCTGTCTGTAATTTGAGAGCCCTATGAATACA	3600
Dp	3653	ACTCAGATATCCACTGTGGCTTTAAAGGCTGTGTGTAATTTGAGAGCCCTATGAATACA	3712
OY	3601	GAAAGGACAAATATTCAAAGTGAACGTGACGGGGCCTAGCTCAACAAGTCTT-----	3651
Dp	3713	GAAAGGACAAATATTCAAAGTGAACGTGACGGGGCCTAGCTCAACAAGTCTT-----	3772
OY	3652	-----CTTGTCTGTGTGTCAGGCA	3669
Dp	3773	CTGATTGACACACACACGCTTACTCTTGACAGCAGAGCTTGTGTGTGTCAGGCA	3832
OY	3670	ATGGCAGTTAATATTTCCGCAAAATGGTTTGGATTGTCATTTGTCAGCTCAAAATGTGTA	3729
Dp	3883	ATGGCAGTTAATATTTCCGCAAAATGGTTTGGATTGTCATTTGTCAGCTCAAAATGTGTA	3892
OY	3730	TATATAATGGAAGGCTTCTGGGTCTTTAGAAAGACGAGATCTATTCAAATCAAGAACCC	3789
Dp	3893	TATATAATGGAAGGCTTCTGGGTCTTTAGAAAGACGAGATCTATTCAAATCAAGAACCC	3952
OY	3790	TTTGAATTTAGATGTTGCTGTATAAAGAAATAAAGTATCTCAATCATATGTGGATTTGAAAT	3849
Dp	3953	TTTGAATTTAGATGTTGCTGTATAAAGAAATAAAGTATCTCAATCATATGTGGATTTGAAAT	4012
OY	3850	GTGTGTGTAACAACCTTTTCCGGCCCGGGTGTGAGATGGCAAGCTCTTATGGAAGTTAAACCTA	3909
Dp	4013	GTGTGTGTAACAACCTTTTCCGGCCCGGGTGTGAGATGGCAAGCTCTTATGGAAGTTAAACCTA	4072
OY	3910	TTAAGTGGCTTATATGTGCTTCAAGAACCAATTTCTCTAGACGAGACATGAGAAAGTGT	3969

Db	4073	TTAAGTGGCTTTATGAGCGCTTCAGAGCAATTTCTGTGACCGAGACGTGAGAAAGTG	4132
Qy	3970	GAATATGATCATGAGAAACCTCAACCTCTATTAGATCTGTAAATGAAACCACTTTTGT	4029
Db	4133	GAATATGATCATGAGAAACCTCAACCTCTATTAGATCTGTAAATGAAACCAAGTTTGT	4192
Qy	4030	GTTAATAATTCCTGCTGTGAGAAACTTTAAAGTTTCAATAACCAAGATGCTTCAGTGTCC	4089
Db	4193	GTTAATAATTCCTGCTGTGAGAAACTTTAAAGTTTCAATAACCAAGATGCTTCAGTGTCC	4252
Qy	4090	ATATGTGATTACTATGAGCCAGGAGACAGCGGTGAAAGTTTCAACTCTGAAAGTGAAG	4149
Db	4253	ATATGTGATTACTATGAGCCAGGAGACAGCGGTGAAAGTTTCAACTCTGAAAGTGAAG	4312
Qy	4150	CTGTCCTCCCTGTGACCTTTTGACGATGATCCAGGGCGCCGTCCTTTTGAGAGATGAGACT	4209
Db	4313	CTGTCCTCCCTGTGACCTTTTGACGATGATCCAGGGCGCCGTCCTTTTGAGAGATGAGACT	4372
Qy	4210	TCAGGCTCCCATCATCACTCTTCAGATATTTTATTTCTGTTCAAGCTTCGTACTTT	4269
Db	4373	TCAGGCTCCCATCATCACTCTTCAGATATTTTATTTCTGTTCAAGCTTCGTACTTT	4432
Qy	4270	ATGGAACCTTGGCGTGTGA	4287
Db	4433	ATGGAACCTTGGCGTGTGA	4450

## RESULT

ID AAD49440 standard; DNA; 5882 BP.

AC AAD49440;

DT 24-MAR-2003 (first entry)

DE Human blood cell surface antigen, CD109 encoding DNA #2.

KW Human; GPI-anchored TGF-beta1 binding protein; r150 protein; cancer;

KW therapy; blood cell surface antigen; CD109; ds.

OS Homo sapiens.

PN W0200285942-A2.

PD 31-OCT-2002.

PF 24-APR-2002; 2002WQ-CA00560

PR 24-APR-2001; 2001US-285713P

XX

XX

XX

XX

PT comprises 150 protein which acts as accessory receptor of TGF-beta,

PT treating cancer -

PS Disclosure; Fig 17; 127pp; English.

CC The invention relates to novel transforming growth factor (TGF)-beta1

CC anchored TGF-beta1 binding protein referred to as r150 which acts as

CC modulating TGF-beta activity, and thus for treating conditions

CC molecules of the invention are used for increasing TGF-beta availability

CC and increase graft success. The present sequence is human blood cell  
CC surface antigen, CD109 encoding DNA. CD109 is a variant of r150 sequence.  
CC Note: This sequence is stated to encode human CD109 protein shown in  
CC figure 17 of the specification, but this does not appear to be the case.

**SQ** Sequence 5882 BP; 1696 A; 1173 C; 1241 G; 1772 T; 0 other;

Query Match	98.5%	Score 4224.4;	DB 25;	Length 5882;
Best Local Similarity	98.8%;	Pred. No. 0;		
Matches 4286; Conservative	0;	Mismatches	1;	Indels 51; Gaps 1;

Qy	1	ATGAGAGGGCCACCGCTCTCTAAGCGCGCCACCTCTCTGCGTGTGACCGCGCGCTG	60
Db	113	ATGCAAGGGCCACCGCTCTCTAAGCGCGCCACCTCTCTGCGTGTGACCGCGCGCTG	172
Qy	61	GCCGTGCTCCCGGGCTCGGTCTTGATGACGCCCGAGGATCATGAGCCCGAGGA	120
Db	173	GCCGTGCTCCCGGGCTCGGTCTTGATGACGCCCGAGGATCATGAGCCCGAGGA	232
Qy	121	AATGGAATATTGGGTGGAGCTTCTGGAACACTGCCCTTACAGGTGACTGTGAAGCG	180
Db	233	AATGGAATATTGGGTGGAGCTTCTGGAACACTGCCCTTACAGGTGACTGTGAAGCG	292
Qy	181	GAGCTGCTCAAGACAGCATCAAACTCTACTGCTCTGTCTGGAAGCAGAAGAGTCTTT	240
Db	293	GAGCTGCTCAAGACAGCATCAAACTCTACTGCTCTGTCTGGAAGCAGAAGAGTCTTT	352
Qy	241	GAAGAAAGGCTCTTTTAAAGACTTACTCTTCCATCACTACTTGAACAGTGCAGATGAG	300
Db	353	GAAGAAAGGCTCTTTTAAAGACTTACTCTTCCATCACTACTTGAACAGTGCAGATGAG	412
Qy	301	ATTATATAGCTACCGTGAACCGAGCTAACCGAGTACCGAGATTTATTTCTTAATGTAC	360
Db	413	ATTATATAGCTACCGTGAACCGAGCTAACCGAGTACCGAGATTTATTTCTTAATGTAC	472
Qy	361	CGCTTATCATTTGAGACCAAGAGATATCTGTCTCATTTCAACAGACAGGCTTATAC	420
Db	473	CGCTTATCATTTGAGACCAAGAGATATCTGTCTCATTTCAACAGACAGGCTTATAC	532
Qy	421	AAGCCAAAGCAAGATGAAGTTTGCATTTGTACACTCTTCTCAGATTTTAAGCTTAC	480
Db	533	AAGCCAAAGCAAGATGAAGTTTGCATTTGTACACTCTTCTCAGATTTTAAGCTTAC	592
Qy	481	AAAACTTTTAAACATTCCTATTAAAGACCCCAATCAAAATTTGATGCCAGGTGTG	540
Db	593	AAAACTTTTAAACATTCCTATTAAAGACCCCAATCAAAATTTGATGCCAGGTGTG	652
Qy	541	TCACAACAAGTGAATCTTGAGATCATTTCCAAACTTTTACGCTATCTTCCATCATTA	600
Db	653	TCACAACAAGTGAATCTTGAGATCATTTCCAAACTTTTACGCTATCTTCCATCATTA	712
Qy	601	CTTGGTGACTGTCTATTCACAGTTCAAGTGAATGACACAGCATATTAATCATTTACG	660
Db	713	CTTGGTGACTGTCTATTCACAGTTCAAGTGAATGACACAGCATATTAATCATTTACG	772
Qy	661	GTTTCAGAAATGATTTACCAAAATTTGAAGACTTTGAGAGCCCTTATATTTGCT	720
Db	773	GTTTCAGAAATGATTTACCAAAATTTGAAGACTTTGAGAGCCCTTATATTTGCT	832
Qy	721	ATGAATTTGAACATTTAAATGATCACTACCGGCAAAATATACATATGGAAGCCAGTG	780
Db	833	ATGAATTTGAACATTTAAATGATCACTACCGGCAAAATATACATATGGAAGCCAGTG	892
Qy	781	AAAGGAGCGTAAACGCTTACATTTTACCCTTATCTCTTTGGGAAAGAGAAAATATTT	840
Db	893	AAAGGAGCGTAAACGCTTACATTTTACCCTTATCTCTTTGGGAAAGAGAAAATATTT	952
Qy	841	ACAAATAACATTTAAGATAAATGATCTGCAAACTCTCTTTAATGATGAAGAGATGAA	900
Db	953	ACAAATAACATTTAAGATAAATGATCTGCAAACTCTCTTTAATGATGAAGAGATGAA	1012
Qy	901	AATGTAATGATTTCTTCAATGACTTTCTGATATCCTGATCTATCTTCCCTGAGACA	960

Db	1013	AATGTAATGCAATTCCTTCAAAATGCACTTTCGAAATACCTGGATCTATCTTCCCTGGACA	1072
Qy	961	GTAAAGAAATTTTAAACCAAGTGAACGAATACGTTTACAGTATTTTCAAGAAATGTAGCACT	1020
Db	1073	GTAGAAATTTTAAACCAAGTGAACGAATTCAGTTTACAGTATTTTCAAGAAATGTAGCACT	1132
Qy	1021	AATGTGTTCTTCAAGCAACATGATTAATCATGATTTTGTATTAATACACTGTCGTG	1080
Db	1133	AATGTGTTCTTCAAGCAACATGATTAATCATGATTTTGTATTAATACACTGTCGTG	1192
Qy	1081	AAGCAATCTCTCAACCTTCAAGCCACTGTGTAAGTAACTGTGCTGATGGCAACCACTG	1140
Db	1193	AAGCAATCTCTCAACCTTCAAGCCACTGTGTAAGTAACTGTGCTGATGGCAACCACTG	1252
Qy	1141	ACTCTTAAGAAAGAAATAATATGTGATATAACATGACACAGAAATCTATATACGAG	1200
Db	1253	ACTCTTAAGAAAGAAATAATATGTGATATAACATGACACAGAAATCTATATACGAG	1312
Qy	1201	TACTGAGCGGATCTTAACAGTGAATAACAGAAAATGAAACCTGTTCAAGAAATTAATAT	1260
Db	1313	TACTGAGCGGATCTTAACAGTGAATAACAGAAAATGAAACCTGTTCAAGAAATTAATAT	1372
Qy	1261	ACTGTCCCCCAAGTGGAACTTTTAAGATTGAATTCCTGAAATCCTGAGATTCACGTAG	1320
Db	1373	ACTGTCCCCCAAGTGGAACTTTTAAGATTGAATTCCTGAAATCCTGAGATTCACGTAG	1432
Qy	1321	CTACAGTTGAAGGCTTATTTCTCTGGTAATGAATAATGACATGACAGTTTCATATGCTGTT	1380
Db	1433	CTACAGTTGAAGGCTTATTTCTCTGGTAATGAATAATGACATGACAGTTTCATATGCTGTT	1492
Qy	1381	AAGTCTCTTAAGATGAACATACATCCAACTAATAAACAAGATGAAATATATAAGTGGGA	1440
Db	1493	AAGTCTCTCTAAGATGAACATACATCCAACTAATAAACAAGATGAAATATATAAGTGGGA	1552
Qy	1441	TCGCTTTTGAAGTTGGTGTGTTAGTGCGCAACAACAGATTGAAGAGATTAACTATATGTA	1500
Db	1553	TCGCTTTTGAAGTTGGTGTGTTAGTGCGCAACAACAGATTGAAGAGATTAACTATATGTA	1612
Qy	1501	GTATCCAGGGGACAGTTGGTGGCGTGAAGAAACAAATTCACAAAGTTCTCTTTAACA	1560
Db	1613	GTATCCAGGGGACAGTTGGTGGCGTGAAGAAACAAATTCACAAAGTTCTCTTTAACA	1672
Qy	1561	CCAGAAATTCCTTGGACTCCCAAAAAGCCTGTGTATTTGTATTAATTTGAAGATGATGGG	1620
Db	1673	CCAGAAATTCCTTGGACTCCCAAAAAGCCTGTGTATTTGTATTAATTTGAAGATGATGGG	1732
Qy	1621	GAAATTTAAGATGATGTTCTPAAAAATTCCTGTTCACGTTGTTTTTAAAAATTAAGATTAAG	1680
Db	1733	GAAATTTAAGATGATGTTCTPAAAAATTCCTGTTCACGTTGTTTTTAAAAATTAAGATTAAG	1792
Qy	1681	CTAATTTGAGTAAAGTGAAAGCTGAAACCATCTGAGAAATCTCTCTTAGAATCTCTGTG	1740
Db	1793	CTAATTTGAGTAAAGTGAAAGCTGAAACCATCTGAGAAATCTCTCTTAGAATCTCTGTG	1852
Qy	1741	ACAAGCCTGACTCCATAGTTGGGATGTGTAGCTGTTACAAAAGTGTGAATCTGATGAAT	1800
Db	1853	ACAAGCCTGACTCCATAGTTGGGATGTGTAGCTGTTACAAAAGTGTGAATCTGATGAAT	1912
Qy	1801	GCTCTTAATGATATTAACAATGAAAAATGTGTCATGATGTTGAACTTTTAACACAGGA	1860
Db	1913	GCTCTTAATGATATTAACAATGAAAAATGTGTCATGATGTTGAACTTTTAACACAGGA	1972
Qy	1861	TATATTTTGAAGCATGTTTCATGAATCTTTTGGCAGTCTTTGAGGAATGTGACCTTGCGTA	1920
Db	1973	TATATTTTGAAGCATGTTTCATGAATCTTTTGGCAGTCTTTGAGGAATGTGACCTTGCGTA	2032
Qy	1921	TTGACAGATCAAAACCTCAAGAAAGATTAATGATGCTGTTTATGACAAATGCAAGATAT	1980
Db	2033	TTGACAGATCAAAACCTCAAGAAAGATTAATGATGCTGTTTATGACAAATGCAAGATAT	2092
Qy	1981	GCTGAGAGGTTTATGAGAGAAATGAAGACATATGTGTAGATATTCATGACTTTTCTTGG	2040
Db	2093	GCTGAGAGGTTTATGAGAGAAATGAAGACATATGTGTAGATATTCATGACTTTTCTTGG	2152

Qy	2041	GGTAGCAGTCCACATGTCGGAAGATTTTCCAGAGACTGGATTGGCTAGACACAC	2100
Db	2153	GGTAGCAGTCCACATGTCGGAAGATTTTCCAGAGACTGGATTGGCTAGACACAC	2212
Qy	2101	ATGGGTTACAGATTTTACCAAGAAATTGAAAGTAACTGTAACCTGATTCTATCACTTCGG	2166
Db	2213	ATGGGTTACAGATTTTACCAAGAAATTGAAAGTAACTGTAACCTGATTCTATCACTTCGG	2272
Qy	2161	GTGGCTACTGGTTTGTGTATCTCTGAGAGACCTGGGCTCTTGACATCAACTACTCCAGTG	2228
Db	2273	GTGGCTACTGGTTTGTGTATCTCTGAGAGACCTGGGCTCTTGACATCAACTACTCCAGTG	2332
Qy	2221	GAGCTCCAGCCTTCCAAACCATTTTTCATTTTTTGAATCTTCCCTACTCTGTTATCCAGA	2280
Db	2333	GAGCTCCAGCCTTCCAAACCATTTTTCATTTTTTGAATCTTCCCTACTCTGTTATCCAGA	2392
Qy	2281	GGTGAAGAAATTGGTTTGGAAATACATATATCCATTTATTTTGAAGAAGTCCAGCTGAGGT	2340
Db	2393	GGTGAAGAAATTGGTTTGGAAATACATATATCCATTTATTTTGAAGAAGTCCAGCTGAGGT	2452
Qy	2341	AAGGTAATCATTTGAGAAAAGTGCACAAATTTGTATCTTAATGACTCTCAAGTGAATAAT	2400
Db	2453	AAGGTAATCATTTGAGAAAAGTGCACAAATTTGTATCTTAATGACTCTCAAAATGAATAAT	2512
Qy	2401	GCCACAGGCCACACAGACAGACCCCTTCTGTTCCCACTGAGAAITGGGGCACTGTTCTTTT	2460
Db	2513	GCCACAGGCCACACAGACAGACCCCTTCTGTTCCCACTGAGAAITGGGGCAACCTGTTCTTTT	2572
Qy	2461	CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCAAGTCAACAGCTCTTCAACCACT	2520
Db	2573	CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCAAGTCAACAGCTCTTCAACCACT	2632
Qy	2521	GCTTCTGATGCTGTCACCCAAGATGATTTTATGTAAGGCTGGAAGATGAAGAAATCATAT	2580
Db	2633	GCTTCTGATGCTGTCACCCAAGATGATTTTATGTAAGGCTGGAAGATGAAGAAATCATAT	2692
Qy	2581	TCACAAATTCATTTTATTTAGACTTGACCTGACACATAGGCTACAGAGTACCCCTGAAAACCTTG	2640
Db	2693	TCACAAATTCATTTTATTTAGACTTGACCTGACACATAGGCTACAGAGTACCCCTGAAAACCTTG	2752
Qy	2641	AGTTTCTCATTTCCCTCTAATACAGTGACTGGCAGTGAAGAAGTTCAATCACTGCAATT	2700
Db	2753	AGTTTCTCATTTCCCTCTAATACAGTGACTGGCAGTGAAGAAGTTCAATCACTGCAATT	2812
Qy	2701	GGAGATGTTCTTGGTCTTTCATCAATGAGCTTAAAGCTTCATTTGCGAGTCCCTTAATGC	2760
Db	2813	GGAGATGTTCTTGGTCTTTCATCAATGAGCTTAAAGCTTCATTTGCGAGTCCCTTAATGC	2872
Qy	2761	TGTGTGTAACAGAAACATGATTAATTTTGGCTCCAAATATTTACATTTTGGATTATCTGACT	2820
Db	2873	TGTGTGTAACAGAAACATGATTAATTTTGGCTCCAAATATTTACATTTTGGATTATCTGACT	2932
Qy	2821	AAAAAGAAACAACCTACAGATTAATTTTGAAGAAAAGAAAGCTCTTTCATTTATGAGCAAGGT	2880
Db	2933	AAAAAGAAACAACCTACAGATTAATTTTGAAGAAAAGAAAGCTCTTTCATTTATGAGCAAGGT	2992
Qy	2881	TACCAAGAGAACTTCTCTATCAGAGGGAAGATGGCTTTTCACTGCTTTTGGGAATAT	2940
Db	2993	TACCAAGAGAACTTCTCTATCAGAGGGAAGATGGCTTTTCACTGCTTTTGGGAATAT	3052
Qy	2941	GACCTTTCTGGAGAGACTTGGTGTGCACGCTTTTGTTTTAAAGTGTTCCTTGAAGCCGAT	3000
Db	3053	GACCTTTCTGGAGAGACTTGGTGTGCACGCTTTTGTTTTAAAGTGTTCCTTGAAGCCGAT	3112
Qy	3001	CCTTAACATAGATTAATGATCAGAAATGTGTTACACAGAACATCACTTGCTTTAAAGACAT	3060
Db	3113	CCTTAACATAGATTAATGATCAGAAATGTGTTACACAGAACATCACTTGCTTTAAAGACAT	3172
Qy	3061	CAGAAATCCAAACGTGAATTTTGGGATCCAGAGAAAGTATTCATATGAGCTTCAAGGT	3120
Db	3173	CAGAAATCCAAACGTGAATTTTGGGATCCAGAGAAAGTATTCATATGAGCTTCAAGGT	3232

[illegible]

Db 4313 CTGCTCTCTGACCTTTGACAGTAGTCACAGGCGCTGCTTGTAGAGATGAGAGCT 4372  
 QY 4210 TCAGGCTCCCATCATCATCTTCAGTCATTTTATTTTTCGTTCAAGCTTCGACTT 4269  
 Db 4373 TCAGGCTCCCATCATCATCTTCAGTCATTTTATTTTTCGTTCAAGCTTCGACTT 4432  
 QY 4270 ATGGAACCTTGGCTGTGA 4287  
 Db 4433 ATGGAACCTTGGCTGTGA 4450  
 RESULT 4  
 ABQ79966 standard; cDNA; 5895 BP.  
 ID ABQ79966  
 XX ABQ79966;  
 AC  
 XX  
 DT 23-DEC-2002 (first entry)  
 XX  
 DE Human CD109 KI-H7 protein encoding cDNA.  
 XX  
 KM CD109; alpha2 macroglobulin; thioester; cerebroprotective; cardiant;  
 KM immunosuppressive; haemostatic; anticoagulant; thrombolytic; human;  
 KM cardiovascular; vasotropic; gene therapy; CD109 KI-H7; gene; ss.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 113..4450  
 FT /tag= a  
 FT /product= "CD109 KI-H7"  
 PN MO200270696-A2.  
 XX 12-SEP-2002.  
 PD 07-MAR-2002; 2002WO-CA00292.  
 PF 07-MAR-2001; 2001US-273814P.  
 PR 07-MAR-2001; 2001US-273814P.  
 XX  
 PA (SCHU/) SCHUH A.  
 PA (SUTH/) SUTHERLAND R D.  
 PI Schub A, Sutherland RD;  
 XX  
 DR WPI; 2002-713450/77.  
 DR P-PSDB; ABB82167.  
 XX  
 PT New CD109 nucleic acids and polypeptides, useful in gene therapy,  
 PT particularly for treating strokes, myocardial infarctions, thrombosis,  
 PT thrombocytopenia, autoimmune diseases, or organ or bone marrow  
 PT transplantation -  
 PT  
 PS Claim 1; Fig 2a; 156pp; English.  
 XX  
 CC The invention relates to isolated nucleic acid molecules encoding CD109  
 CC polypeptides. These nucleic acid molecules include the human CDNA  
 CC sequences comprising CD109 KI, CD109 KI-H7, CD109 K15 or their variants.  
 CC CD109 is a novel member of the alpha2 macroglobulin (alpha2M)/C3, C4, C5  
 CC family of thioester-containing proteins. The CD109 polypeptides can be  
 CC expressed by standard recombinant methodology. The CD109 nucleic acid,  
 CC CD109 polypeptide or its mimetic, or the anti-CD109 antibody is useful  
 CC for treating or detecting a disease or disorder, e.g. conditions  
 CC associated with endothelial activation, platelet activation, activation  
 CC of the coagulation or fibrinolytic systems, activation of T lymphocytes  
 CC and of the complement system, quantitative or qualitative abnormalities  
 CC of platelet function, increased or impaired platelet aggregation and/or  
 CC activation, increased or impaired activation of the coagulation and/or  
 CC fibrinolytic systems, or impaired or increased immune activation. These  
 CC are also useful for treating cardiovascular disorders, stroke, myocardial  
 CC infarction, thrombosis, embolism, peripheral vascular disease,  
 CC thrombocytopenia, thrombocythemia, autoimmune diseases, organ

CC transplantation, or bone marrow transplantation. The CD109 nucleic acids  
 CC or their antisense nucleotide sequence are useful in gene therapy for  
 CC treating these conditions. The present sequence represents the human  
 CC CD109 KI-H7 cDNA sequence.  
 XX  
 SQ Sequence 5895 BP; 1709 A; 1173 C; 1241 G; 1772 T; 0 other;  
 Query Match 98.5%; Score 4224.4; DB 24; Length 5895;  
 Best Local Similarity 98.8%; Pred. No. 0;  
 Matches 4286; Conservative 0; Mismatches 1; Indels 51; Gaps 1;  
 QY 1 ATGCAAGGCGCCACCGCTCTGACCGCGCCCACTCTCTGCGTGTGACCGCGCGCTG 60  
 Db 113 ATGCAAGGCGCCACCGCTCTGACCGCGCCCACTCTCTGCGTGTGACCGCGCGCTG 172  
 QY 61 GCCGTGGCTCCCGGGCGCTGGTTTCTGGTGAACGCCCGAGGATCATGAGCCGGAGGA 120  
 Db 173 GCCGTGGCTCCCGGGCGCTGGTTTCTGGTGAACGCCCGAGGATCATGAGCCGGAGGA 232  
 QY 121 AATGTACTAATTTGGGATGAGCTTCTGGAACTGCGCTTCAAGTGAAGTGAAGCG 180  
 Db 233 AATGTACTAATTTGGGATGAGCTTCTGGAACTGCGCTTCAAGTGAAGTGAAGCG 292  
 QY 181 GAGCTGCTCAAGACAGCATCAAACTCACTGTCTCTGTCTGTGAAGCAAGAGTCTTT 240  
 Db 293 GAGCTGCTCAAGACAGCATCAAACTCACTGTCTCTGTCTGTGAAGCAAGAGTCTTT 352  
 QY 241 GAAAAAGGCTCTTTTAAACACTTACTCTTCTCATCACTCTGAAACAGTGCAGATGAG 300  
 Db 353 GAAAAAGGCTCTTTTAAACACTTACTCTTCTCATCACTCTGAAACAGTGCAGATGAG 412  
 QY 301 ATTTATGAGCTACGTTGTAACGCGAGTACCGCAGATGAGATTTTATTCTTAATGATAC 360  
 Db 413 ATTTATGAGCTACGTTGTAACGCGAGTACCGCAGATGAGATTTTATTCTTAATGATAC 472  
 QY 361 CGCTTATCATTTAGACCAAGAAATATCTGTCTTCAATCAAGACAAAGGCTTATAC 420  
 Db 473 CGCTTATCATTTAGACCAAGAAATATCTGTCTTCAATCAAGACAAAGGCTTATAC 532  
 QY 421 AAGCCAAAGCAAGATGAAAGTTTGGCATTTGTACACTCTTCTCAAGATTTTAAAGCTTAC 480  
 Db 533 AAGCCAAAGCAAGATGAAAGTTTGGCATTTGTACACTCTTCTCAAGATTTTAAAGCTTAC 592  
 QY 481 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAATCAATTTGATGATCAACAGTGTG 540  
 Db 593 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAATCAATTTGATGATCAACAGTGTG 652  
 QY 541 TCACAAACAAGTGATCTTGAGTCAATTTCCAAAATTTTTCAGTATCTTCCATCCAAATA 600  
 Db 653 TCACAAACAAGTGATCTTGAGTCAATTTCCAAAATTTTTCAGTATCTTCCATCCAAATA 712  
 QY 601 CTGTGTGATGCTGTATTTCAAGTTCAAGTGAATGACACAGATATTTATCAATTTAG 660  
 Db 713 CTGTGTGATGCTGTATTTCAAGTTCAAGTGAATGACACAGATATTTATCAATTTAG 772  
 QY 661 GTTTCAGATATGATATCAAAAATTTGAATGATTTTGACAGACCACTTATTTGTTCT 720  
 Db 773 GTTTCAGATATGATATCAAAAATTTGAATGATTTTGACAGACCACTTATTTGTTCT 832  
 QY 721 ATGAATTTCAAGCATTTAAATGATACATCAACGCAAGTATTAATATGAGGAGGAGTGTG 780  
 Db 833 ATGAATTTCAAGCATTTAAATGATACATCAACGCAAGTATTAATATGAGGAGGAGTGTG 892  
 QY 781 AAGGAGACGTAGAGCTTACATTTTACCTTTATCTTTGGGGAAGAGAAAATATTT 840  
 Db 893 AAGGAGACGTAGAGCTTACATTTTACCTTTATCTTTGGGGAAGAGAAAATATTT 952  
 QY 841 ACAAATAATTAAGATATAAGATCTGCAAACTTCTTTAATGAGGAGAGTGA 900  
 Db 953 ACAAATAATTAAGATATAAGATCTGCAAACTTCTTTAATGAGGAGAGTGA 1012  
 QY 901 AATGTAAATGATTTCTTCAATGAGCTTCTGAAATACCTGATCTATCTTCCCTGAGACA 960

Db 1013 AATGTAATGATCTTCTTAATGAGCTTTCTGAAATCCTGGAATCTAATCTTCCCTGGAGCA 1072  
Qy 961 GTAGAAATTTTAAACACAGTGAACAGATCAGTTACAGGTATTTTCAAGAAATGTAAGCACT 1020  
Db 1073 GTAGAAATTTTAAACACAGTGAACAGATCAGTTACAGGTATTTTCAAGAAATGTAAGCACT 1132  
Qy 1021 AATGTTCTTCAAGCAACATGATTAATCAATGAGTTTGTATTAATTAATTAATTAATTAAT 1080  
Db 1133 AATGTTCTTCAAGCAACATGATTAATCAATGAGTTTGTATTAATTAATTAATTAATTAAT 1192  
Qy 1081 AAGCATCTCTCACTTCAAGCAACATGATTAATCAATGAGTTTGTATTAATTAATTAATTAAT 1140  
Db 1193 AAGCATCTCTCACTTCAAGCAACATGATTAATCAATGAGTTTGTATTAATTAATTAATTAAT 1252  
Qy 1141 ACTCTTGAAGAAAGAAAGAAATTAATGATTAATCAATGAGTTTGTATTAATTAATTAATTAAT 1200  
Db 1253 ACTCTTGAAGAAAGAAAGAAATTAATGATTAATCAATGAGTTTGTATTAATTAATTAATTAAT 1312  
Qy 1201 TACTGAGCGGATCTTAACAGTGAATCAGAAATGAGAACTGTTCAAGAAATTAATTAAT 1260  
Db 1313 TACTGAGCGGATCTTAACAGTGAATCAGAAATGAGAACTGTTCAAGAAATTAATTAAT 1372  
Qy 1261 ACTGCCCCCAAGTGAACCTTTTAAGTTGAATTCCTGAGAGATTCAGAGATTCAGAGAT 1320  
Db 1373 ACTGCCCCCAAGTGAACCTTTTAAGTTGAATTCCTGAGAGATTCAGAGATTCAGAGAT 1432  
Qy 1321 CTACAGTTGAAGGCTTATTTCTTGGTAGTAAAGTAGACATGGAGTTCAATGCTGTT 1380  
Db 1433 CTACAGTTGAAGGCTTATTTCTTGGTAGTAAAGTAGACATGGAGTTCAATGCTGTT 1492  
Qy 1381 AAGTCTCTAGTAAGACATACATCCAACTAAACAAAGATGAAATTAATTAATTAATTAAT 1440  
Db 1493 AAGTCTCTAGTAAGACATACATCCAACTAAACAAAGATGAAATTAATTAATTAATTAAT 1552  
Qy 1441 TCGCTTTTGAAGTGTGTAGTGAACAAAGATTAAGAGATTAAGTAAATTAATTAATTAAT 1500  
Db 1553 TCGCTTTTGAAGTGTGTAGTGAACAAAGATTAAGAGATTAAGTAAATTAATTAATTAAT 1612  
Qy 1501 GTATTCAGAGGAGACGTTGGTGTAGAGAAACAAATTAATTAATTAATTAATTAATTAAT 1560  
Db 1613 GTATTCAGAGGAGACGTTGGTGTAGAGAAACAAATTAATTAATTAATTAATTAATTAAT 1672  
Qy 1561 CCAGAAATTTCTGAGCTCCAAAGCCTGTGTAATTTGATTAATTAATTAATTAATTAAT 1620  
Db 1673 CCAGAAATTTCTGAGCTCCAAAGCCTGTGTAATTTGATTAATTAATTAATTAATTAAT 1732  
Qy 1621 GAAATTAATTAATTAATTAATTTCTGCTGAGCTGTTTAAATTAATTAATTAATTAAT 1680  
Db 1733 GAAATTAATTAATTAATTAATTTCTGCTGAGCTGTTTAAATTAATTAATTAATTAAT 1792  
Qy 1681 CTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1740  
Db 1793 CTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1852  
Qy 1741 ACACAGCTGATCTCATAGTTGGATTTGATGCTTTGCAAAAGTGTGAATCTGATGAT 1800  
Db 1853 ACACAGCTGATCTCATAGTTGGATTTGATGCTTTGCAAAAGTGTGAATCTGATGAT 1912  
Qy 1801 GCTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1860  
Db 1913 GCTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1972  
Qy 1861 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1920  
Db 1973 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2032  
Qy 1921 TTGACAGATGCAACCTCAGCAAGATTAATTAATTAATTAATTAATTAATTAATTAAT 1980  
Db 2033 TTGACAGATGCAACCTCAGCAAGATTAATTAATTAATTAATTAATTAATTAATTAAT 2092  
Qy 1981 GCTGAGAGTTTATGAGAGAAATGAAGACATTAATTAATTAATTAATTAATTAATTAAT 2040  
Db 2093 GCTGAGAGTTTATGAGAGAAATGAAGACATTAATTAATTAATTAATTAATTAATTAAT 2152

Qy 2041 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGCTTGGATTTGGCTAGACACCAAC 2100  
Db 2153 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGCTTGGATTTGGCTAGACACCAAC 2212  
Qy 2101 ATGGGTTTACAGATTTTCCAGAAATTTGAAGTAATCTGTACTGATTTTATCACTTCTTGG 2160  
Db 2213 ATGGGTTTACAGATTTTCCAGAAATTTGAAGTAATCTGTACTGATTTTATCACTTCTTGG 2272  
Qy 2161 GTGGCTATCTGGTTTGTGATCTGTGAGGACCTGGGCTTGGAGTACCACTACTCCAGTG 2220  
Db 2273 GTGGCTATCTGGTTTGTGATCTGTGAGGACCTGGGCTTGGAGTACCACTACTCCAGTG 2332  
Qy 2221 GAGCTCCAAAGCTTCCAAACCATTTTCAATTTTGTGAATCTTCCCTAATCTGTATCAGA 2280  
Db 2333 GAGCTCCAAAGCTTCCAAACCATTTTCAATTTTGTGAATCTTCCCTAATCTGTATCAGA 2392  
Qy 2281 GGTGAAGATTTTGGTTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2340  
Db 2393 GGTGAAGATTTTGGTTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2452  
Qy 2341 AAGGTATCAATGAGAAAGTGAACAAATTTGATTAATTAATTAATTAATTAATTAAT 2400  
Db 2453 AAGGTATCAATGAGAAAGTGAACAAATTTGATTAATTAATTAATTAATTAATTAAT 2512  
Qy 2401 GCCACAGGCGCACAGAGACCTTCTGTTCCAGTGAAGATGGGCACTGTTCTTTT 2460  
Db 2513 GCCACAGGCGCACAGAGACCTTCTGTTCCAGTGAAGATGGGCACTGTTCTTTT 2572  
Qy 2461 CCATCAAGGCGCAACATCTGGAGAAATTTCTATCACTCACTCACTCTTCAACCACT 2520  
Db 2573 CCATCAAGGCGCAACATCTGGAGAAATTTCTATCACTCACTCACTCTTCAACCACT 2632  
Qy 2521 GCTTGTGATGCTGTACCCGATGATTTTGAAGAGCTGAAGAAATGAAGAAATCAAT 2580  
Db 2633 GCTTGTGATGCTGTACCCGATGATTTTGAAGAGCTGAAGAAATGAAGAAATCAAT 2692  
Qy 2581 TCACAAATCCATCTTATTAAGACTGACGACAAATAGCTCAGAGTACCTGAAACCTTGG 2640  
Db 2693 TCACAAATCCATCTTATTAAGACTGACGACAAATAGCTCAGAGTACCTGAAACCTTGG 2752  
Qy 2641 AGTTTCTCATTTCTCTTAATTAATGAGTCTGCAAGTGAAGATTAATTAATTAATTAAT 2700  
Db 2753 AGTTTCTCATTTCTCTTAATTAATGAGTCTGCAAGTGAAGATTAATTAATTAATTAAT 2812  
Qy 2701 GGAAGTCTTGTGCTTCTTCCATCAATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2760  
Db 2813 GGAAGTCTTGTGCTTCTTCCATCAATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2872  
Qy 2761 TGTGTGAACAGAAATGATTAATTTTGTCCAAATTAATTAATTAATTAATTAATTAAT 2820  
Db 2873 TGTGTGAACAGAAATGATTAATTTTGTCCAAATTAATTAATTAATTAATTAATTAAT 2932  
Qy 2821 AAAAAGAAACATGACAGATTAATTTGAAAGAAAGCTTCTTCAATTAATTAATTAATTAAT 2880  
Db 2933 AAAAAGAAACATGACAGATTAATTTGAAAGAAAGCTTCTTCAATTAATTAATTAATTAAT 2992  
Qy 2881 TACCAAGAGAACTTCTCTATCAGAGGAAAGATGCTCTTTCAGTCTTTTGGGAATTAAT 2940  
Db 2993 TACCAAGAGAACTTCTCTATCAGAGGAAAGATGCTCTTTCAGTCTTTTGGGAATTAAT 3052  
Qy 2941 GACCTTCTGGAGAGACTTGGTGTGAGCTTTTGTATTAAGATTTTCTTGAAGCCGAT 3000  
Db 3053 GACCTTCTGGAGAGACTTGGTGTGAGCTTTTGTATTAAGATTTTCTTGAAGCCGAT 3112  
Qy 3001 CTTTACATAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3060  
Db 3113 CTTTACATAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3172  
Qy 3061 CAGAAATCCAAAGGTTTGGATTTGGATCCAGGAAGATTAATTAATTAATTAATTAATTAAT 3120  
Db 3173 CAGAAATCCAAAGGTTTGGATTTGGATCCAGGAAGATTAATTAATTAATTAATTAATTAAT 3232

OY	3121	GGCAATAAATCCAGTAACCTTACAGCCCTATATGTGAATCTTCCTCCGGAATATAGA	3180
Db	3223	GGCAATAAATCCAGTAACCTTACAGCCCTATATGTGAATCTTCCTCCGGAATATAGA	3292
OY	3181	AAGTATACGCTTAACTTATGTGCAGAGCTCTATCCATTTTTTGGAGTCGAATTGAGT	3240
Db	3293	AAGTATACGCTTAACTTATGTGCAGAGCTCTATCCATTTTTTGGAGTCGAATTGAGT	3352
OY	3241	AGAGGAATTTCAAGCAATTATACCTAGCCCTTATTAATTTCGATTTGTCATCACTGCGG	3300
Db	3353	AGAGGAATTTCAAGCAATTATACCTAGCCCTTATTAATTTCGATTTGTCATCACTGCGG	3412
OY	3301	AGCTCTAAAGGGAAGGAAGCTTTGAATATGCTGACTTGGAGAGAGAAACAAGAGTGGC	3360
Db	3413	AGCTCTAAAGGGAAGGAAGCTTTGAATATGCTGACTTGGAGAGAGAAACAAGAGTGGC	3472
OY	3361	ATGCAATTTCTGGGTGCATCTAGAGTCCAACTTTCTGACTCTGGCAGCCAGCTCCCTG	3420
Db	3473	ATGCAATTTCTGGGTGCATCTAGAGTCCAACTTTCTGACTCTGGCAGCCAGCTCCCTG	3532
OY	3421	GATATTTGAAGTTGCAAGCTTATGCACTGCTCTACACTTCTTACAATTTCAAGCTTTGAG	3480
Db	3533	GATATTTGAAGTTGCAAGCTTATGCACTGCTCTACACTTCTTACAATTTCAAGCTTTGAG	3592
OY	3481	GGAATCCCAATTATAGAGGTGGCTTAAGCAAGAAATATGCTGGGTGGTTTTCGACT	3540
Db	3593	GGAATCCCAATTATAGAGGTGGCTTAAGCAAGAAATATGCTGGGTGGTTTTCGACT	3652
OY	3541	ACTCAGAGTACCACTGTGGCTTTTAAAGGCTGTCTGTAATTTGAGCCCTATGAATACA	3600
Db	3653	ACTCAGAGTACCACTGTGGCTTTTAAAGGCTGTCTGTAATTTGAGCCCTATGAATACA	3712
OY	3601	GAAAGCAAAATATCCAAAGTGAACCGTGAACGGGCGTTAGCTCACCAGTCTT-----	3651
Db	3713	GAAAGCAAAATATCCAAAGTGAACCGTGAACGGGCGTTAGCTCACCAGTCTT-----	3772
OY	3652	-----CTTGGCTGTGTATCAGCCA	3669
Db	3773	CTGATTTGACACACAACAACCGCTTACTCCCTTCAGACAGAGACTGTGGTGTATCACA	3832
OY	3670	ATGACAGTAAATATTTCCGAAATAGTTTGGATTTTGCTATTTGTCACGCTCAATGTGTGA	3729
Db	3883	ATGACAGTAAATATTTCCGAAATAGTTTGGATTTTGCTATTTGTCACGCTCAATGTGTGA	3892
OY	3730	TATATGTGAAGGCTTCTGGTCTTCTTGAAGAAGCAAGATCTATCCAAATCAAGAAGCC	3789
Db	3893	TATATGTGAAGGCTTCTGGTCTTCTTGAAGAAGCAAGATCTATCCAAATCAAGAAGCC	3952
OY	3790	TTTGATTTAGATGTTGCTGTATAAAGAAATATGAATGATCTCAATCATGTGATTTGAAT	3849
Db	3953	TTTGATTTAGATGTTGCTGTATAAAGAAATATGAATGATCTCAATCATGTGATTTGAAT	4012
OY	3850	GTGTGTACAACCTTTTGGGCGCCGGGTATGAGTGGCAATGGCTCTTATGGAAGTTAACCTA	3909
Db	4013	GTGTGTACAACCTTTTGGGCGCCGGGTATGAGTGGCAATGGCTCTTATGGAAGTTAACCTA	4072
OY	3910	TTAAGTGGCTTTATGTGTCCTTACAGAACATTTCTCTGAGCCGAGACAGTGAAGAAAGTG	3969
Db	4073	TTAAGTGGCTTTATGTGTCCTTACAGAACATTTCTCTGAGCCGAGACAGTGAAGAAAGTG	4132
OY	3970	GAATATGATCATGGAATACTCAACCTCTATTATTAATTTCTGTAATGAACCCAGTTTTGT	4029
Db	4133	GAATATGATCATGGAATACTCAACCTCTATTATTAATTTCTGTAATGAACCCAGTTTTGT	4192
OY	4030	GTTTAATATTCCTGCGTGTGAGAAACCTTTAAATTTCAATATCCCAAGATGCTTCAGTTC	4089
Db	4193	GTTTAATATTCCTGCGTGTGAGAAACCTTTAAATTTCAATATCCCAAGATGCTTCAGTTC	4252
OY	4090	ATAAGTGATTTACTATGAGCCAAAGAGACAGCGGTGAGAGTTTCAACTCTGAAAGTGAAG	4149
Db	4253	ATAAGTGATTTACTATGAGCCAAAGAGACAGCGGTGAGAGATTTCAACTCTGAAAGTGAAG	4312
OY	4150	CTGTCTCTCTGTACCTTTGCACTGATGTCCAGGAGTCCGCTCTCTGTGAGATGTGAGCT	4209

Db	4313	CTGTCCTCCTGTGAACCTTTGCAGCGATGATCCAGGGCTGCCCTCTTGTGAGGATGGACT	
Oy	4210	TCAGGCTCCCATCACTCTTCAGTCATTTTTATTTCTGTTTCAAGCTTGACTTT	
Dd	4373	TCAGGCTCCCATCACTCTTCAGTCATTTTTATTTCTGTTTCAAGCTTGACTTT	
Oy	4270	ATGGAACCTTGGCTGTGA 4287	
Db	4433	ATGGAACCTTGGCTGTGA 4450	
 RESULT 5 AAD49435 standard; DNA; 5883 BP.			
ID	AAD49435		
XX	AAD49435;		
XX			
XX	24-MAR-2003 (first entry)		
DE	Human r150 DNA #1.		
XX			
KW	Human; GPI-anchored TGF-beta1 binding protein; r150 protein; cancer; glycosylphosphatidylinositol; transforming growth factor-beta1; therapy; gene; ds.		
OS	Homo sapiens.		
PH			
FT	Key	Location/Qualifiers	
FT	CDS	113..4450	
FT		/tag= a	
FT		/product= "Human r150 protein #2"	
FT		/trans_except= (pos:2219..2221, aa:Xaa)	
FN		/note= "Xaa corresponds to Ser, Tyr"	
XX			
PN	WO200285942-A2.		
PD	31-OCT-2002.		
PF	24-APR-2002; 2002MO-CA00560.		
PR	24-APR-2001; 2001US-285713P.		
PA	14-FEB-2002; 2002US-356163P.		
XX	(UYMC-) UNIV MCGILL.		
PI	Philip A, Tam B;		
DR	WPI; 2003-093100/08.		
DR	P-PSTDB; NAEJ2013.		
PT	Novel transforming growth factor (TGF)-beta 1 binding reagent which comprises r150 protein which acts as accessory receptor of TGF-beta, useful for negatively modulating TGF-beta activity, and thus for treating cancer -		
PS	Claim 7; Page 100-103; 127gp; English.		
XX			
XX	The invention relates to novel transforming growth factor (TGF) beta1 binding reagent which comprises a glycosylphosphatidyl inositol (GPI) - anchored TGF-beta1 binding protein referred to as r150 which acts as an accessory receptor of TGF-beta. The invention is used for negatively modulating TGF-beta activity, and thus for treating conditions characterised by overproduction of TGF-beta, such as cancer. Antisense molecules of the invention are used for increasing TGF-beta availability and increase graft success. The present sequence is human r150 protein encoding DNA.		
SO	Sequence 5883 BP; 1696 A; 1173 C; 1241 G; 1772 T; 1 other;		
Query Match	98.5%; Score 4223.4; DB 25; Length 5883;		
Best Local Similarity	98.8%; Pred. No. 0;		
Matches 4285; Conservative 0; Mismatches 2; Indels 51; Gaps 1;			



QY 1 ATGAGGGCCCAACGCTCTGACCGCCGACCTCTCTGCTGCTGCAACCGCGCTG 60  
DB 113 ATGACGGGCCACCGCTCTGACCGCGCCGACCTCTCTGCTGCTGCAACCGCGCTG 172  
QY 61 GCCGTGGCTCCCGGGCTCGGCTTCTGCTGACAGCGCCGAGGATCATAGCCCGAGGA 120  
DB 173 GCCGTGGCTCCCGGGCTCGGCTTCTGCTGACAGCGCCGAGGATCATAGCCCGAGGA 232  
QY 121 AATGACTATTTGGGGTGGAGCTTCTGGACACTGCGCTTCAAGGTGACTGTAGAGCG 180  
DB 233 AATGACTATTTGGGGTGGAGCTTCTGGACACTGCGCTTCAAGGTGACTGTAGAGCG 292  
QY 181 GAGCTGCTCAAGACAGCATCAAACTCACTGCTCTGCTCTGGAAGCAGAGAGTCTTT 240  
DB 293 GAGCTGCTCAAGACAGCATCAAACTCACTGCTCTGCTCTGGAAGCAGAGAGTCTTT 352  
QY 241 GAAAAAGGCTCTTTTAAAGACCTTACTCTTCCATCATCTTCAAGAGTGCAGATGAG 300  
DB 353 GAAAAAGGCTCTTTTAAAGACCTTACTCTTCCATCATCTTCAAGAGTGCAGATGAG 412  
QY 301 ATTATGAGCTACGCTGTACCGGACGTACCGAGATGAGATTTTATCTCTAATAGTACC 360  
DB 413 ATTATGAGCTACGCTGTACCGGACGTACCGAGATGAGATTTTATCTCTAATAGTACC 472  
QY 361 CGCTTATCATTTGAGACCAAGAGAAATATCTGCTTCAATTCAAGACAGAGGCTTTATAC 420  
DB 473 CGCTTATCATTTGAGACCAAGAGAAATATCTGCTTCAATTCAAGACAGAGGCTTTATAC 532  
QY 421 AAGCCAAAGCAAGAGTGAAGTTCGCACTGTGTAACCTCTTCAAGATTTTAAAGCTTTAC 480  
DB 533 AAGCCAAAGCAAGAGTGAAGTTCGCACTGTGTAACCTCTTCAAGATTTTAAAGCTTTAC 592  
QY 481 AAAACCTCTTAAACATCTCTAATTAAGGACCCCAATTAATTTGATCCAGTGTG 540  
DB 593 AAAACCTCTTAAACATCTCTAATTAAGGACCCCAATTAATTTGATCCAGTGTG 652  
QY 541 TCACAAACAAAGTGAATCTTGAGTCAATTTCCAAAATTTTCACTATCTTCCATCCATA 600  
DB 653 TCACAAACAAAGTGAATCTTGAGTCAATTTCCAAAATTTTCACTATCTTCCATCCATA 712  
QY 601 CTGCTGACGTGCTTATCAAGTTCAGTGAATGACAGACATATTAATCAATTTTCAAG 660  
DB 713 CTGCTGACGTGCTTATCAAGTTCAGTGAATGACAGACATATTAATCAATTTTCAAG 772  
QY 661 GTTTCAGAAATATGATTAACCAAAATTTGAAGTGAATTTGACAGACCAATTAATTTTCT 720  
DB 773 GTTTCAGAAATATGATTAACCAAAATTTGAAGTGAATTTGACAGACCAATTAATTTTCT 832  
QY 721 ATGAATCTAAGCAATTTAAATGATACATCAAGCAAAAGTATATAGTGAAGCCAGTG 780  
DB 833 ATGAATCTAAGCAATTTAAATGATACATCAAGCAAAAGTATATAGTGAAGCCAGTG 892  
QY 781 AAAGAGAGCGTACGCTTACATTTTATCTTTTATCTTTTGGGAAAAGAAAATATTT 840  
DB 893 AAAGAGAGCGTACGCTTACATTTTATCTTTTATCTTTTGGGAAAAGAAAATATTT 952  
QY 841 ACAAACAACTTAAATGATTAAGTATGATCTGCAAACTTCTTTTAAATGATGAAGATGAAA 900  
DB 953 ACAAACAACTTAAATGATTAAGTATGATCTGCAAACTTCTTTTAAATGATGAAGATGAAA 1012  
QY 901 AATGTAATGATCTTCAATAGACCTTCTGAATACCTGATCTATCTTCCCTGAGACA 960  
DB 1013 AATGTAATGATCTTCAATAGACCTTCTGAATACCTGATCTATCTTCCCTGAGACA 1072  
QY 961 GTAGAAATTTTAAACAGTACAGAAATCACTTACAGTATTTTCAAGAAATGTAAGACT 1020  
DB 1073 GTAGAAATTTTAAACAGTACAGAAATCACTTACAGTATTTTCAAGAAATGTAAGACT 1132  
QY 1021 AATGTGTTCTTCAAGCAATGATTAATCATTTGAGTTTTTGAATTAATCACTGCTG 1080  
DB 1133 AATGTGTTCTTCAAGCAATGATTAATCATTTGAGTTTTTGAATTAATCACTGCTG 1192

QY 1081 AAGCATCTCTCAACTTCAAGCCACTGTGAAGATACTGCTGTAGTGAACCAACTG 1140  
DB 1193 AAGCATCTCTCAACTTCAAGCCACTGTGAAGATACTGCTGTAGTGAACCAACTG 1252  
QY 1141 ACTTTGAAAGAAAGAAATTAATGATGATTAAGTGAACAGTGAACAGAGAAATTAATGAG 1200  
DB 1253 ACTTTGAAAGAAAGAAATTAATGATGATTAAGTGAACAGTGAACAGAGAAATTAATGAG 1312  
QY 1201 TACTGAGCGGATCTAAACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1260  
DB 1313 TACTGAGCGGATCTAAACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1372  
QY 1261 ACTGTCCCAAGAGTGAAGCTTTTAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1320  
DB 1373 ACTGTCCCAAGAGTGAAGCTTTTAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1432  
QY 1321 CTACAGTTGAAGGCTTATTTCTTGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1380  
DB 1433 CTACAGTTGAAGGCTTATTTCTTGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1492  
QY 1381 AAGTCTCTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1440  
DB 1493 AAGTCTCTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1552  
QY 1441 TCGCTTTTGAAGTGTGTTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1500  
DB 1553 TCGCTTTTGAAGTGTGTTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1612  
QY 1501 GTATCCAGGGGACAGTGTGCTGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1560  
DB 1613 GTATCCAGGGGACAGTGTGCTGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1672  
QY 1561 CCAAGAAATCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1620  
DB 1673 CCAAGAAATCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1732  
QY 1621 GAAATTAATGATGATCTTAAATTTCTGTTCAAGTGTGTTTAAATTAATGATTAAG 1680  
DB 1733 GAAATTAATGATGATGATCTTAAATTTCTGTTCAAGTGTGTTTAAATTAATGATTAAG 1792  
QY 1681 CTATATTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1740  
DB 1793 CTATATTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1852  
QY 1741 ACAAGCTGACCTCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1800  
DB 1853 ACAAGCTGACCTCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1912  
QY 1801 GCCTTAATGATATTAATCAATGAGAAATGTGTCATGATGATGAGTGAACCTTTATTAACAGAGA 1860  
DB 1913 GCCTTAATGATATTAATCAATGAGAAATGTGTCATGATGATGAGTGAACCTTTATTAACAGAGA 1972  
QY 1861 TATTTATTTAGGCAATGTTCAATGAATCTTTTGCAGTCTTTTCAAGAAATGTGAGTCTGAGTA 1920  
DB 1973 TATTTATTTAGGCAATGTTCAATGAATCTTTTGCAGTCTTTTCAAGAAATGTGAGTCTGAGTA 2032  
QY 1921 TTGACAGATGCAAACTCAAGAGTGAATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1980  
DB 2033 TTGACAGATGCAAACTCAAGAGTGAATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 2092  
QY 1981 GCTGAGAGGTTTATGAGAGAAATGAAGACATATTTGATGATTAATGATGATGATGATGATGATGATGATGAT 2040  
DB 2093 GCTGAGAGGTTTATGAGAGAAATGAAGACATATTTGATGATTAATGATGATGATGATGATGATGATGATGATGAT 2152  
QY 2153 GTAGAGAGTCAATGATGCGAAGGCAATTTTCCAGAGTGTGATTTGGCTGAGACACCAAC 2212  
DB 2101 ATGGTTTACAGATTTTAAAGAAATTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2160  
QY 2213 ATGGTTTACAGATTTTAAAGAAATTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2272  
DB 2161 GTGGCTACTGCTTGTGATCTGTGAGGACTGAGTCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2220



Db 2273 GTGGCTACGCTGTTGTGATCTCTGAGGACCTGGGCTTGGACATCACTCCAGAG 2332  
 QY 2221 GAGCTCCAGACCTTCCACCATTTTCATTTTTTGAATCTTCCCTACCTGTATACAA 2280  
 Db 2333 GAGCTCCAGACCTTCCACCATTTTCATTTTTTGAATCTTCCCTACCTGTATACAA 2392  
 QY 2281 GGTGAAGAAATTTGCTTGGAAATTAACATATTTCAATTATTTGAAGATCCACTGAGTT 2340  
 Db 2393 GGTGAAGAAATTTGCTTGGAAATTAACATATTTCAATTATTTGAAGATCCACTGAGTT 2452  
 QY 2341 AAGGTATATTTGAGAAAAGTGAACAATTTGATATTTCAATGACTTCAAGTAATTAAT 2400  
 Db 2453 AAGGTATATTTGAGAAAAGTGAACAATTTGATATTTCAATGACTTCAATGAAATTAAT 2512  
 QY 2401 GCCACAGGCGACAGAGACCCCTTGTGTCCAGAGAGATGGGCGAATCTTCTTTT 2460  
 Db 2513 GCCACAGGCGACAGAGACCCCTTGTGTCCAGAGAGATGGGCGAATCTTCTTTT 2572  
 QY 2461 CCCATCAGGCGAACACATCTGGAGAAATTCATACAGTCAAGCTCTTCAACCACT 2520  
 Db 2573 CCCATCAGGCGAACACATCTGGAGAAATTCATACAGTCAAGCTCTTCAACCACT 2632  
 QY 2521 GCTTCTGATGCTGTCCACAGATGATTTTAAGGCTGAAGGAAATGAAATATCATAT 2580  
 Db 2633 GCTTCTGATGCTGTCCACAGATGATTTTAAGGCTGAAGGAAATGAAATATCATAT 2692  
 QY 2581 TCACATCATCTTATTTAGACTGACATAGCTACAGATACCTCTGAAACTTTG 2640  
 Db 2693 TCACATCATCTTATTTAGACTGACATAGCTACAGATACCTCTGAAACTTTG 2752  
 QY 2641 AGTTTCTCATTTCTCTCTAATACAGTACGTCGACAGTGAAGATTCAGATCACTGCAAT 2700  
 Db 2753 AGTTTCTCATTTCTCTCTAATACAGTACGTCGACAGTGAAGATTCAGATCACTGCAAT 2812  
 QY 2701 GGAGATGTTCTTGTCTTCCATCAATGAGCTTACCTCATGATTTGATGAGCTTATGAGC 2760  
 Db 2813 GGAGATGTTCTTGTCTTCCATCAATGAGCTTACCTCATGATTTGATGAGCTTATGAGC 2872  
 QY 2761 TGTGTGAACAGAACTGATTAATTTTGTCCAAATTTTACATTTTGAATTTATCTGACT 2820  
 Db 2873 TGTGTGAACAGAACTGATTAATTTTGTCCAAATTTTACATTTTGAATTTATCTGACT 2932  
 QY 2821 AAAAAAGAAACAACCTGACAGATATTTGAAGAAAAAGCTCTTCAATTATGAGCGAAGT 2880  
 Db 2933 AAAAAAGAAACAACCTGACAGATATTTGAAGAAAAAGCTCTTCAATTATGAGCGAAGT 2992  
 QY 2881 TACAGAGAGAACTTCTCTATCAGAGGAGAAATGCTCTTCAAGTCTTTTGGAAATTAAT 2940  
 Db 2993 TACAGAGAGAACTTCTCTATCAGAGGAGAAATGCTCTTCAAGTCTTTTGGAAATTAAT 3052  
 QY 2941 GACCCCTTCTGGAGACCTTGGTGTGACGTTTGTTTAAGATGTTCTTGAAGCCGAT 3000  
 Db 3053 GACCCCTTCTGGAGACCTTGGTGTGACGTTTGTTTAAGATGTTCTTGAAGCCGAT 3112  
 QY 3001 CCTTACATGATATTTGATCAGATGATGTTTACACAGAACATACCTTGGCTTAAAGGACAT 3060  
 Db 3113 CCTTACATGATATTTGATCAGATGATGTTTACACAGAACATACCTTGGCTTAAAGGACAT 3172  
 QY 3061 CAGAAATCCAAACGCTGAATTTTGGATCCAGAGAGAGATTCATAGTAGCTTCAAGGT 3120  
 Db 3173 CAGAAATCCAAACGCTGAATTTTGGATCCAGAGAGAGATTCATAGTAGCTTCAAGGT 3232  
 QY 3121 GGCATTAATAAGTCCAGTAACTTACAGCTTATTTGATCTTCTCTGGATATTAAG 3180  
 Db 3233 GGCATTAATAAGTCCAGTAACTTACAGCTTATTTGATCTTCTCTGGATATTAAG 3292  
 QY 3181 AAGTATCAGCTTAACATGATGAGAGAGTATATCTTTTGGAGCTGAATTCAGT 3240  
 Db 3293 AAGTATCAGCTTAACATGATGAGAGAGTATATCTTTTGGAGCTGAATTCAGT 3352  
 QY 3241 AAGAGATTTCAAGCAATTAATCTAGCCCTTAATCTTATGATGATCAGTGGGG 3300

Db 3353 AAGAGAAATTTCAACAATTAATCTTAGCCCTTAATTAATGATTTGATCAGTGGGG 3412  
 QY 3301 AGTCTTAAGAGAGAGAGATTTTGAATATGCTGACTTGAAGAGAGAACAAAGAGTGGC 3360  
 Db 3413 AGTCTTAAGAGAGAGAGATTTTGAATATGCTGACTTGAAGAGAGAACAAAGAGTGGC 3472  
 QY 3361 ATGCAATTCGGGTGTCAATCAGAGTCCAACTTTCTGACTCTGGCAGCAGCTCCCTG 3420  
 Db 3473 ATGCAATTCGGGTGTCAATCAGAGTCCAACTTTCTGACTCTGGCAGCAGCTCCCTG 3532  
 QY 3421 GATATTAAGATTTGACAGCTTATGACCTGCTCAACTTTTCAATTTTCAACTTTGAG 3480  
 Db 3533 GATATTAAGATTTGACAGCTTATGACCTGCTCAACTTTTCAATTTTCAACTTTGAG 3592  
 QY 3481 GGAATCCCAATTAAGAGTGGCTTAAGCAGGAGAAATTAACCTGGGTGGTTTGCACT 3540  
 Db 3593 GGAATCCCAATTAAGAGTGGCTTAAGCAGGAGAAATTAACCTGGGTGGTTTGCACT 3652  
 QY 3541 ACTCAGATTAACCACTGTGGCTTTAAAGGCTGTCTGAAATTTGACAGCCCTAATGATACA 3600  
 Db 3653 ACTCAGATTAACCACTGTGGCTTTAAAGGCTGTGTGAAATTTGACAGCCCTAATGATACA 3712  
 QY 3601 GAAAGCAAAATATCCAAAGTACCGGTGACGGGGCTTAGCTCAACAAGTCT- 3651  
 Db 3713 GAAAGCAAAATATCCAAAGTACCGGTGACGGGGCTTAGCTCAACAAGTCT- 3772  
 QY 3652 -----CTTGTGTGTGACAGCA 3669  
 Db 3773 CTGATTAACACACACACACACCGCTTACTCTTCAACAGAGAGCTTGTGTGACAGCA 3832  
 QY 3670 ATGCAATTAATTAATTTCCGCAAAATGTTTGAATTTGATTTGTCAGCTCAATGTTGA 3729  
 Db 3833 ATGCAATTAATTAATTTCCGCAAAATGTTTGAATTTGTCAGCTCAATGTTGA 3892  
 QY 3730 TATAATGTAAGGCTTGTGGGTCTTTCAGAAAGCAGAAATATTCAAATCAAGAACCC 3789  
 Db 3893 TATAATGTAAGGCTTGTGGGTCTTTCAGAAAGCAGAAATATTCAAATCAAGAACCC 3952  
 QY 3790 TTTGATTTAAGATGTTGCTGTAAAGAAATTAAGATGATTCATCAATGATGATTTGAT 3849  
 Db 3953 TTTGATTTAAGATGTTGCTGTAAAGAAATTAAGATGATTCATCAATGATGATTTGAT 4012  
 QY 3850 GTGTGTCAACGCTTTTGGGCGCGGTGAGAGTGGCATGAGCTTATGAAAGTTAACTTA 3909  
 Db 4013 GTGTGTCAACGCTTTTGGGCGCGGTGAGAGTGGCATGAGCTTATGAAAGTTAACTTA 4072  
 QY 3910 TTAAGTGGCTTTATGAGTCCCTTCAGAAAGCAATTTCTGACGAGACAGTGAAGAAATG 3969  
 Db 4073 TTAAGTGGCTTTATGAGTCCCTTCAGAAAGCAATTTCTGACGAGACAGTGAAGAAATG 4132  
 QY 3970 GAATATGATCATGAGAAACTCAACTCTATTTAATTTCTGTAATTTGAAGAACCAAGTTTGT 4029  
 Db 4133 GAATATGATCATGAGAAACTCAACTCTATTTAATTTCTGTAATTTGAAGAACCAAGTTTGT 4192  
 QY 4030 GTTAATTAATTTCTGCTGTGAGAAACTTTAAAGTTTCAATTAACCAAGATGCTTCACTGTC 4089  
 Db 4133 GTTAATTAATTTCTGCTGTGAGAAACTTTAAAGTTTCAATTAACCAAGATGCTTCACTGTC 4252  
 QY 4030 ATAGTGAATTAATTAAGCAGAGAGACAGGCGGTGAGAAATTAACCTTGAAGTGAAG 4149  
 Db 4253 ATAGTGAATTAATTAAGCAGAGAGACAGGCGGTGAGAAATTAACCTTGAAGTGAAG 4312  
 QY 4150 CTGTCTCTCTGTGACCTTTGCAAGTATTCAGAGGCTGCGCTCTTGTGAGAGATGAGCT 4209  
 Db 4313 CTGTCTCTCTGTGACCTTTGCAAGTATTCAGAGGCTGCGCTCTTGTGAGAGATGAGCT 4372  
 QY 4210 TCAGGCTGCCAATCACTCTTCAAGTATTTTATTTTCTGTGTTCAAGCTCTGACTT 4269  
 Db 4373 TCAGGCTGCCAATCACTCTTCAAGTATTTTATTTTCTGTGTTCAAGCTCTGACTT 4432  
 QY 4270 ATGGAATTTGGCTGTGA 4287  
 Db 4433 ATGGAATTTGGCTGTGA 4450

RESULT 6  
AB079965

ID ABQ79965 standard; cDNA; 4761 BP.

AC ABQ79965;

DT 23-DEC-2002 (first entry)

Human CD109 K1 variant protein encoding cDNA.

KM CD109; alpha2 macroglobulin; thioester; cerebroprotective; cardiant;  
KM immunosuppressive; haemostatic; anticoagulant; thrombolytic; human;  
KM cardiovascular; vasotropic; gene therapy; CD109 KI; variant; gene; ss

**Homo sapiens.**

FH	Key	Location/Qualifiers
FT	CDS	113..4450

```

/ product= "CD109 K1 variant"
FT

```

PN WO200270696-A2.

PD 12-SEP-2002

PF 07-MAR-2002; 2002WO-CA00292.

PR 07-MAR-2001; 2001US-273814P.

PA (SCHU/) SCHUH A.

XX

XX

P-PSDB; ABB82166.

PT New CD109 nucleic

PT particularly for treating strokes, myocardial infarctions, thrombosis  
PT thrombocytopenia, autoimmune diseases, or organ or bone marrow  
PT transplantation -

PS Claim 1; Fig 1b; 156pp; English.

The invention relates to isolated nucleic acid molecules encoding CD109 polypeptides. These nucleic acid molecules include the human cDNA sequences comprising CD109 K1, CD109 K1-H7, CD109 K15 or their variants. CD109 is a novel member of the alpha2 macroglobulin (alpha2M)/C3, C4, C5 family of thioester-containing proteins. The CD109 polypeptides can be expressed by standard recombinant methodology. The CD109 nucleic acid, CD109 polypeptide or its mimetic, or the anti-CD109 antibody is useful for treating or detecting a disease or disorder, e.g. conditions associated with endothelial activation, platelet activation, activation of the coagulation or fibrinolytic systems, activation of T lymphocytes and of the complement system, quantitative or qualitative abnormalities of platelet function, increased or impaired platelet aggregation and activation, increased or impaired activation of the coagulation and/or fibrinolytic systems, or impaired or increased immune activation. These are also useful for treating cardiovascular disorders, stroke, myocardial infarction, thrombosis, embolism, peripheral vascular disease, thrombocytopenia, thrombocythemia, autoimmune diseases, organ transplantation, or bone marrow transplantation. The CD109 nucleic acids or their antisense nucleotide sequence are useful in gene therapy for treating these conditions. The present sequence represents the human CD109 K1 variant cDNA sequence.

Sequence 4761 BP; 1419 A; 913 C; 1016 G; 1413 T; 0 other;

Query Match	98.5%;	Score 4222.8;	DB 24;	Length 4761;
Best Local Similarity	98.8%;	Pred. NO. 0;		
Matches 4285; Conservative	0;	Mismatches 2;	Indels 51;	Gaps 1.

OY		1	TTGCAAGGGCCACCGCTCTGTGAACCGCGGCCACACTCTCTGCAGTGTSCAACGCGCGCCTG	60
Db		113	ATGCAAGGGCCACCGCTCTGTGAACCGCGGCCACACTCTCTGCAGTGTSCAACGCGCGCCTG	172
OY		61	GCCGTGCTCCCGGGCTCCGGTTTTCTGTGTACAGCCCCAAGGATCATCAGGCCCGAGGA	120
Db		173	GCCGTGCTCCCGGGCTCCGGTTTTCTGTGTACAGCCCCAAGGATCATCAGGCCCGAGGA	232
OY		121	AATGTGACTATATGGGGTGAGAGCTCTGTGAAACAATGGCCCTCACAGGTGACTGTGAAGGGC	180
Db		233	AATGTGACTATATGGGGTGAGAGCTCTGTGAAACAATGGCCCTCACAGGTGACTGTGAAGGGC	292
OY		181	GAGCTGCTCAAGACAGCATCAAACCTTCACGTCTGTCTCTGTGAGCAGAAGAGTCTTT	240
Db		293	GAGCTGCTCAAGACAGCATCAAACCTTCACGTCTGTCTCTGTGAGCAGAAGAGTCTTT	352
OY		241	GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACTCTGTGAACATGTGCAGATGAG	300
Db		353	GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACTCTGTGAACATGTGCAGATGAG	412
OY		301	ATTATATGAGCTACGTGTACCCGGAACATACCAGATGAGATTTTATCTCTAATATGATACC	360
Db		413	ATTATATGAGCTACGTGTACCCGGAACATACCAGATGAGATTTTATCTCTAATATGATACC	472
OY		361	CGCTTATCATTTGAGACCAAGAGAAATATCTGTCTTCATTCAAAACAGACAGAGCCCTTATAC	420
Db		473	CGCTTATCATTTGAGACCAAGAGAAATATCTGTCTTCATTCAAAACAGACAGAGCCCTTATAC	532
OY		421	AAGCCAAGCAGAAGAGGAGTTTGGATTTGTAACAATCTTCCAGATTTTAAAGCCTTAC	480
Db		533	AAGCCAAGCAGAAGAGGAGTTTGGATTTGTAACAATCTTCCAGATTTTAAAGCCTTAC	592
OY		481	AAAACCTCTTTAAACATCTTCATTAAAGAGCCCAATAATTGATCCAAAGTGGTGG	540
Db		593	AAAACCTCTTTAAACATCTTCATTAAAGAGCCCAATAATTGATCCAAAGTGGTGG	652
OY		541	TCACAACAAAGTGATCTTTGAGATCTTTCCAAAACCTTTCAAGCTATCTTCCCATCCATA	600
Db		653	TCACAACAAAGTGATCTTTGAGATCTTTCCAAAACCTTTCAAGCTATCTTCCCATCCATA	712
OY		601	CTTGTGATCAGGTGTATTAACAAGTTCAAGTAATGACAGACATATTAATCAATCATTTTCAAG	660
Db		713	CTTGTGATCAGGTGTATTAACAAGTTCAAGTAATGACAGACATATTAATCAATCATTTTCAAG	772
OY		661	GTITTCAGATATATGATTATACCAAAATTTGAAGTAGCTTTGCAGACACCATTAATATGTTCT	720
Db		773	GTITTCAGATATATGATTATACCAAAATTTGAAGTAGCTTTGCAGACACCATTAATATGTTCT	832
OY		721	ATGAATCTTAAGCATTTAANTGTATGACATACCGGCAAAAGTATCATATGTGGGAAGCCAGTG	780
Db		833	ATGAATCTTAAGCATTTAANTGTATGACATACCGGCAAAAGTATCATATGTGGGAAGCCAGTG	892
OY		781	AAAGGAGACGTAAACGTTACATTTTAACTCTTTATCCTTTTGGGGAAAAGAAAGAAAATATTT	840
Db		893	AAAGGAGACGTAAACGTTACATTTTAACTCTTTATCCTTTTGGGGAAAAGAAAGAAAATATTT	952
OY		841	ACAAAAACATTTAAGATTAATGTGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA	900
Db		953	ACAAAAACATTTAAGATTAATGTGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA	1012
OY		901	AATGTATNGATCTTCAAAATGAGCTTTGGAATPACTGGATCTATCTTCCCTGGAGCA	960
Db		1013	AATGTATNGATCTTCAAAATGAGCTTTGGAATPACTGGATCTATCTTCCCTGGAGCA	1072
OY		961	GTAGAAAATTTTAAACCAAGTGAAGATCAAGTACAGSTATTTTCAAGAAAATGTATAGACT	1020
Db		1073	GTAGAAAATTTTAAACCAAGTGAAGATCAAGTACAGSTATTTTCAAGAAAATGTATAGACT	1132
OY		1021	AATGTGTTCTTCAAGCAACATGATATCATATGAGTTTGTGATTAATACTATCTGTCTTG	1080
Db		1133	AATGTGTTCTTCAAGCAACATGATATCATATGAGTTTGTGATTAATACTATCTGTCTTG	1192



```

Db 3353 AGAGGATTTGACACATTAATCTAGCCCTTAATTAATGATGTCATCAGTGGGG 3412
Qy 3301 AGTCTTAAGGAGGAGGAGCTTTGATATATGCTGACTTGGAGAGGAGCAAGAGTGGC 3360
Db 3413 AGTCTTAAGGAGGAGGAGCTTTGATATATGCTGACTTGGAGAGGAGCAAGAGTGGC 3472
Qy 3361 ATGCATTTCTGGGTCTCATCAGAGTCCAACTTTCTGACTCTGGAGCCAGCTCCCTG 3420
Db 3473 ATGCATTTCTGGGTCTCATCAGAGTCCAACTTTCTGACTCTGGAGCCAGCTCCCTG 3532
Qy 3421 GATATGAGTGGAGGCTTAATGAGCACTGCTCAGCACTTCAATTAATTAATTAATTAAT 3480
Db 3533 GATATGAGTGGAGGCTTAATGAGCACTGCTCAGCACTTCAATTAATTAATTAATTAAT 3592
Qy 3481 GGAATCCCAATTAATGAGTGGCTAAGCAGGCAAGAAATAGCTGGGTGGTTCATCT 3540
Db 3593 GGAATCCCAATTAATGAGTGGCTAAGCAGGCAAGAAATAGCTGGGTGGTTCATCT 3652
Qy 3541 ACTCAGATACCACTGTGGCTTTTAAAGGCTGTCTGTAATTTGAGCCCTTAATTAATTAAT 3600
Db 3653 ACTCAGATACCACTGTGGCTTTTAAAGGCTGTCTGTAATTTGAGCCCTTAATTAATTAAT 3712
Qy 3601 GAAAGGACAAATATTCAGAGTGCAGCGGCGCTTACTCACCAGTCTCT----- 3651
Db 3713 GAAAGGACAAATATTCAGAGTGCAGCGGCGCTTACTCACCAGTCTCTCTTAAGTTT 3772
Qy 3652 -----CTTGTGTGTGTGAGCA 3669
Db 3773 CTGATTTGACACACACACCGCTTACTCTTACAGACAGAGCTTGTGTGTGTGTGTGTGT 3832
Qy 3670 ATGGCAGTTAATTTTCCGCAATGTTTGAATTTGCTAATTTGTCACTCAATGTTGTA 3729
Db 3833 ATGGCAGTTAATTTTCCGCAATGTTTGAATTTGCTAATTTGTCACTCAATGTTGTA 3892
Qy 3730 TATATGAGAGGCTTCTGGGTCTTCTGAAGAAGAAATCTATTCACAAATCAAGAAAGCC 3789
Db 3893 TATATGAGAGGCTTCTGGGTCTTCTGAAGAAGAAATCTATTCACAAATCAAGAAAGCC 3952
Qy 3790 TTTGATTTAGATGTTGCTGTAAAGAAATTAAGATGATCTCAATCATGTGATTTGTAAT 3849
Db 3953 TTTGATTTAGATGTTGCTGTAAAGAAATTAAGATGATCTCAATCATGTGATTTGTAAT 4012
Qy 3850 GTGTGTACAGCTTTCCGGGCCCGGTGAGAGTGCATGCTCTTAATGAAATTAACCTTA 3909
Db 4013 GTGTGTACAGCTTTCCGGGCCCGGTGAGAGTGCATGCTCTTAATGAAATTAACCTTA 4072
Qy 3910 TTAAGTGGCTTTATGAGTGCCTTTCAGAGCAATTTCTGAGCGGAGCAGTGAAGAAAGTG 3969
Db 4073 TTAAGTGGCTTTATGAGTGCCTTTCAGAGCAATTTCTGAGCGGAGCAGTGAAGAAAGTG 4132
Qy 3970 GAATATGATCAGTGAAGAACTCAACCTTAATTAATGATCTGTAAATGAAGAACCAAGTTTGT 4029
Db 4133 GAATATGATCAGTGAAGAACTCAACCTTAATTAATGATCTGTAAATGAAGAACCAAGTTTGT 4192
Qy 4030 GTTAATATCTGCTGTGAGAACTTTAAAGTTTCAATTAATCCAAAGTCTTCAAGTCTC 4089
Db 4193 GTTAATATCTGCTGTGAGAACTTTAAAGTTTCAATTAATCCAAAGTCTTCAAGTCTC 4252
Qy 4090 ATATGAGTATCTATGAGCCAGAGACAGCGGCTGAGAAATTAATCTGAAAGTGAAG 4149
Db 4253 ATATGAGTATCTATGAGCCAGAGACAGCGGCTGAGAAATTAATCTGAAAGTGAAG 4312
Qy 4150 CTGTCTCTCTGAGCCTTTGAGAGTGTCCAGGGGCTGCTCTTGTGAGAGTGAAGT 4269
Db 4313 CTGTCTCTCTGAGCCTTTGAGAGTGTCCAGGGGCTGCTCTTGTGAGAGTGAAGT 4372
Qy 4210 TCAGGCTCCCATCATCACTCTTCACTGATTTTATTTCTGTTCAAGCTTCTGACTTT 4269
Db 4373 TCAGGCTCCCATCATCACTCTTCACTGATTTTATTTCTGTTCAAGCTTCTGACTTT 4432
Qy 4270 ATGGAACCTTTGGCTGTGA 4287
Db 4433 ATGGAACCTTTGGCTGTGA 4450

```

```

RESULT 7
ABQ79967
ID ABQ79967 standard; cDNA; 5895 BP.
XX
AC ABQ79967;
XX
DT 23-DEC-2002 (first entry)
XX
DE Human CD109 KI-H7 variant protein encoding cDNA.
XX
KW CD109; alpha2 macroglobulin; thioester; cerebroprotective; cardiant;
KW immunosuppressive; haemostatic; anticoagulant; thrombolytic; human;
KW cardiovascular; vasotropic; gene therapy; CD109 KI-H7; variant; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 113..4450
FT /tag= a
FT /product= "CD109 KI-H7 variant"
XX
EN W0200270696-A2.
XX
PD 12-SEP-2002.
XX
PF 07-MAR-2002; 2002M0-CA00292.
XX
PR 07-MAR-2001; 2001US-273814P.
XX
PA (SCHU/) SCHUH A.
PA (SUTH/) SUTHERLAND R D.
XX
PI Schuh A, Sutherland RD;
XX
DR WPI; 2002-713450/77.
DR P-PSDB; ABB82168.
XX
PT New CD109 nucleic acids and polypeptides, useful in gene therapy,
PT particularly for treating strokes, myocardial infarctions, thrombosis,
PT thrombocytopenia, autoimmune diseases, or organ or bone marrow
PT transplantation -
XX
PS Claim 1; Fig 2b; 156bp; English.
XX
CC The invention relates to isolated nucleic acid molecules encoding CD109
CC polypeptides. These nucleic acid molecules include the human cDNA
CC sequences comprising CD109 KI, CD109 KI-H7, CD109 KI5 or their variants.
CC CD109 is a novel member of the alpha2 macroglobulin (alpha2M)/C3, C4, C5
CC family of thioester-containing proteins. The CD109 polypeptides can be
CC expressed by standard recombinant methodology. The CD109 nucleic acid,
CC CD109 polypeptide or its mimetic, or the anti-CD109 antibody is useful
CC for treating or detecting a disease or disorder, e.g. conditions
CC associated with endothelial activation, platelet activation, activation
CC of the coagulation and fibrinolytic systems, activation of T lymphocytes
CC and of the complement system, quantitative or qualitative abnormalities
CC of platelet function, increased or impaired platelet aggregation and
CC activation, increased or impaired activation of the coagulation and/or
CC fibrinolytic systems, or impaired or increased immune activation. These
CC are also useful for treating cardiovascular disorders, stroke, myocardial
CC infarction, thrombosis, embolism, peripheral vascular disease,
CC thrombocytopenia, thrombocytopenia, autoimmune diseases, organ
CC transplantation, or bone marrow transplantation. The CD109 nucleic acids
CC or their antisense nucleotide sequence are useful in gene therapy for
CC treating these conditions. The present sequence represents the human
CC CD109 KI-H7 variant cDNA sequence.
XX
SQ Sequence 5895 BP; 1708 A; 1174 C; 1241 G; 1772 T; 0 other;

```

```

Query Match 98.5%; Score 4222.8; DB 24; Length 5895;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 4285; Conservative 0; Mismatches 2; Indels 51; Gaps 1;

```

QY 1 ATGAGGGCCACCGCTCTGACCGCGCCACCTCTCTGCGTGTGACCGCGCGCTG 60  
Db 113 ATGAGGGGCCACCGCTCTGACCGCGCCACCTCTCTGCGTGTGACCGCGCGCTG 172  
QY 61 GCCGTGGCTCCCGGGCTCGGTTTCTGTGTGACGCGCCCGAGGATCATCGGCCGAGGA 120  
Db 173 GCCGTGGCTCCCGGGCTCGGTTTCTGTGTGACGCGCCCGAGGATCATCGGCCGAGGA 232  
QY 121 AATGTGACTATTGGGGTGGAGCTTCTGTGAACACTGCGCCCTTCA CAGGTGACTGTGAAGGCG 180  
Db 223 AATGTGACTATTGGGGTGGAGCTTCTGTGAACACTGCGCCCTTCA CAGGTGACTGTGAAGGCG 292  
QY 181 GAGCTGCTCAAGACAGATCAAACTCTGCTCTGTCTGTGAAGCAGAAAGAGTCTTT 240  
Db 293 GAGCTGCTCAAGACAGATCAAACTCTGCTCTGTCTGTGAAGCAGAAAGAGTCTTT 352  
QY 241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCATCACTACCTGTGAACAGTGCAGATGAG 300  
Db 353 GAAAAAGGCTCTTTTAAGACACTTACTCTTCATCACTACCTGTGAACAGTGCAGATGAG 412  
QY 301 ATTATGAGCTAGTGTAAACGGAGTACCGAGATCCAGAGATGAGATTATTCCTTAATGATACC 360  
Db 413 ATTATGAGCTAGTGTAAACGGAGTACCGAGATCCAGAGATGAGATTATTCCTTAATGATACC 472  
QY 361 CGCTTATCATTTGAGACCAAGAGAAATATCTGTCTTCAATCAACAGACAGAGGCTTTATAC 420  
Db 473 CGCTTATCATTTGAGACCAAGAGAAATATCTGTCTTCAATCAACAGACAGAGGCTTTATAC 532  
QY 421 AAGCCAAAGCAAGAGTGAAGTTTGGCATTTGTTACACTCTTCTCAAGATTTTAAGCCTTAC 480  
Db 533 AAGCCAAAGCAAGAGTGAAGTTTGGCATTTGTTACACTCTTCTCAAGATTTTAAGCCTTAC 592  
QY 481 AAAACCTCTTTAAACATCTCATTAAGACCCCAATCAAAATTGATCCAAAGGCTTG 540  
Db 593 AAAACCTCTTTAAACATCTCATTAAGACCCCAATCAAAATTGATCCAAAGGCTTG 652  
QY 541 TCACAAACAAAGTATCTTGAAGTCAATTTCCAAAACCTTTCAAGTATCTTCCATCCAAAT 600  
Db 653 TCACAAACAAAGTATCTTGAAGTCAATTTCCAAAACCTTTCAAGTATCTTCCATCCAAAT 712  
QY 601 CTGGTGAATGTGTATTCAGTCAAGTGAATGACAGACCAATATTCATCATTTGACG 660  
Db 713 CTGGTGAATGTGTATTCAGTCAAGTGAATGACAGACCAATATTCATCATTTGACG 772  
QY 661 GTTTTCAAGATATGTATTAACAAAATTTGAAGTCACTTTGACAGACCAATATTCATCATTTG 720  
Db 773 GTTTTCAAGATATGTATTAACAAAATTTGAAGTCACTTTGACAGACCAATATTCATCATTTG 832  
QY 721 ATGAATTCCTAAGCATTTAAATGTATCAATCAAGGCAAAAGTATATCATTTGGAAGGCAAGT 780  
Db 833 ATGAATTCCTAAGCATTTAAATGTATCAATCAAGGCAAAAGTATATCATTTGGAAGGCAAGT 892  
QY 781 AAAAGAGACGTAAACGCTTACATTTTACCTTTATCTTTTGGGAAGAAAGAAATATTT 840  
Db 893 AAAAGAGACGTAAACGCTTACATTTTACCTTTATCTTTTGGGAAGAAAGAAATATTT 952  
QY 841 ACAAAGACATTTAAATATTAATGTATCTGTGAACCTTCTCTTTTAATGAAGAAAGATGA 900  
Db 953 ACAAAGACATTTAAATATTAATGTATCTGTGAACCTTCTCTTTTAATGAAGAAAGATGA 1012  
QY 901 AATGTAAATGATTTCTTCAAAATGACCTTTCTGAATACCTGATCTATCTTCCCTGAGACA 960  
Db 1013 AATGTAAATGATTTCTTCAAAATGACCTTTCTGAATACCTGATCTATCTTCCCTGAGACA 1072  
QY 961 GTAGAAATTTTAAACAAGTGA CAGAAATCAGTATCAAGATTTTCAAGAAATGTAAGCACT 1020  
Db 1073 GTAGAAATTTTAAACAAGTGA CAGAAATCAGTATCAAGATTTTCAAGAAATGTAAGCACT 1132  
QY 1021 AATGTGTTCTTCAAGACAGATATTCATCATTTGAGTTTGTGATTAATCTACTGCTTG 1080  
Db 1133 AATGTGTTCTTCAAGACAGATATTCATCATTTGAGTTTGTGATTAATCTACTGCTTG 1192

QY 1081 AAGCATCTCTCAACTTTCACAGCACCTGTGAAGTAACTGTGTGATGGAACCAACTG 1140  
Db 1193 AAGCATCTCTCAACTTTCACAGCACCTGTGAAGTAACTGTGTGATGGAACCAACTG 1252  
QY 1141 ACTCTTGAAGAAAGAAATTAATGTAGTCACTAACAGTGAACACAGAAACTATCTAG 1200  
Db 1253 ACTCTTGAAGAAAGAAATTAATGTAGTCACTAACAGTGAACACAGAAACTATCTAG 1312  
QY 1201 TACTGAGACGGATCTAACAGTGAAGAAATGAGAAATGGAAGCTGTCAGAAAAATAATAT 1260  
Db 1313 TACTGAGACGGATCTAACAGTGAAGAAATGAGAAATGGAAGCTGTCAGAAAAATAATAT 1372  
QY 1261 ACTGTCCCCCAAAGTGAACCTTTTAAGTAAATTCCTGAGAGATTCAGTGAAG 1320  
Db 1373 ACTGTCCCCCAAAGTGAACCTTTTAAGTAAATTCCTGAGAGATTCAGTGAAG 1432  
QY 1321 CTACAGTTGAAGGCTATTTCTTGTGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1380  
Db 1433 CTACAGTTGAAGGCTATTTCTTGTGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1492  
QY 1381 AAGTCTCTAGTAAAGACATACATCACTAAACCAAGAGATGAAGAAATTAAGGTGGA 1440  
Db 1493 AAGTCTCTAGTAAAGACATACATCACTAAACCAAGAGATGAAGAAATTAAGGTGGA 1552  
QY 1441 TCGCTTTTGAAGTGTGTGTAGTGAAGCAAGATTTGAAGAGTGAAGTGAAGTGAAGTGA 1500  
Db 1553 TCGCTTTTGAAGTGTGTGTAGTGAAGCAAGATTTGAAGAGTGAAGTGAAGTGAAGTGA 1612  
QY 1501 GTATCAGGGGACAGTGTGTGTGTAGTGAAGCAAGATTTGAAGAGTGAAGTGAAGTGAAG 1560  
Db 1613 GTATCAGGGGACAGTGTGTGTGTAGTGAAGCAAGATTTGAAGAGTGAAGTGAAGTGAAG 1672  
QY 1561 CGAGAAATTTCTGAGCTCCAAAAGCCGTGTAAATGTGTATTAATGAAGTGAAGTGAAG 1620  
Db 1673 CGAGAAATTTCTGAGCTCCAAAAGCCGTGTAAATGTGTATTAATGAAGTGAAGTGAAG 1732  
QY 1621 GAAATTTAATGATGATGTTCTAAAAATTCCTGTTCAAGTGTGTTTAAAAATGAAGTGAAG 1680  
Db 1733 GAAATTTAATGATGATGTTCTAAAAATTCCTGTTCAAGTGTGTTTAAAAATGAAGTGAAG 1792  
QY 1681 CTATTTGAGATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1740  
Db 1793 CTATTTGAGATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1852  
QY 1741 ACACAGCTGATCTCATGTTGGGATGTGATCTGTGAACAAAGTGAAGTGAAGTGAAGTGA 1800  
Db 1853 ACACAGCTGATCTCATGTTGGGATGTGATCTGTGAACAAAGTGAAGTGAAGTGAAGTGA 1912  
QY 1801 GCCTCTAATGATTAATTAACAATGGAAGTGTGATGAGTGAAGTGAAGTGAAGTGAAGTGA 1860  
Db 1913 GCCTCTAATGATTAATTAACAATGGAAGTGTGATGAGTGAAGTGAAGTGAAGTGAAGTGA 1972  
QY 1861 TATTAATTTAGGATGTTTCAATGAATCTTTTCAAGTCTTTCAGAAATGTGAAGTGAAGTGA 1920  
Db 1973 TATTAATTTAGGATGTTTCAATGAATCTTTTCAAGTCTTTCAGAAATGTGAAGTGAAGTGA 2032  
QY 1921 TTGACAGATGCAAACTTCAAGAGATTAATGTAGTGTGTTTGAACAAAGTGAAGTGAAGTGA 1980  
Db 2033 TTGACAGATGCAAACTTCAAGAGATTAATGTAGTGTGTTTGAACAAAGTGAAGTGAAGTGA 2092  
QY 1981 GCTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2040  
Db 2093 GCTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2152  
QY 2041 GGTAGCACTCAATGTCGAAAGCATTTTTCAGAGACTTGGATTTGGCTAGACCAAC 2100  
Db 2153 GGTAGCACTCAATGTCGAAAGCATTTTTCAGAGACTTGGATTTGGCTAGACCAAC 2212  
QY 2101 ATGGGTTTACAGATTTTACAGAAATTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2160  
Db 2213 ATGGGTTTACAGATTTTACAGAAATTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2272  
QY 2161 GTGGCTACTGGTTTGTGATCTGTAGAGACCTGGGCTTGTGAAGTGAAGTGAAGTGAAGTGA 2220

2273 GTGGCTACTGGTGTGGATCTCTGAGGACCTGGGCTTGGACATACAACTACCTCAGTG 2332  
2221 GAGCTCCAAAGCCTTCCAAACATTTTTCATTTTGTGAATCTCCCTACTCTGTATCAGA 2280  
2333 GAGCTCCAAAGCCTTCCAAACATTTTTCATTTTGTGAATCTCCCTACTCTGTATCAGA 2392  
2281 GGTGAAGAATTTGCTTGGAAATACTATATTCATATTTTGAAGAATGCCACTGAGGT 2340  
2293 GGTGAAGAATTTGCTTGGAAATACTATATTCATATTTTGAAGAATGCCACTGAGGT 2452  
2341 AAGTAATCATTTGAAAAAGTGAACAAATTTGATATTTATGACTTCAAGTGAATTAAT 2400  
2453 AAGTAATCATTTGAAAAAGTGAACAAATTTGATATTTATGACTTCAAGTGAATTAAT 2512  
2401 GCCACAGGCCACACAGACACCTTCTGGTCTCCAGTGAAGTGGGGCACTGTTCTTTT 2460  
2513 GCCACAGGCCACACAGACACCTTCTGGTCTCCAGTGAAGTGGGGCACTGTTCTTTT 2572  
2461 CCCATCAGGCCACACATCTGGAGAAAATTCCTATCACAAGTCAAGCTTTTCAACCACT 2520  
2573 CCCATCAGGCCACACATCTGGAGAAAATTCCTATCACAAGTCAAGCTTTTCAACCACT 2632  
2521 GCTTCTGATGCTGTACCCCAATGATTTTATGAAAAGCTGAAGATGAAAAATCATAT 2580  
2633 GCTTCTGATGCTGTACCCCAATGATTTTATGAAAAGCTGAAGATGAAAAATCATAT 2692  
2581 TCACAATCATCTTATTTAGACTGTGACAAATGGCTGACAGAGTACCCCTGAAAACTTTG 2640  
2693 TCACAATCATCTTATTTAGACTGTGACAAATGGCTGACAGAGTACCCCTGAAAACTTTG 2752  
2641 AGTTTCTCATTTCTCTCTAATACAGTGACTGCGACAGTGAAGAGTTCAGATCATGCAAT 2700  
2753 AGTTTCTCATTTCTCTCTAATACAGTGACTGCGACAGTGAAGAGTTCAGATCATGCAAT 2812  
2701 GGAATGTTCTTGTGCTTCCATCAATGCTTAGCTTACCTTATGATTTGGATGCTTATGCT 2760  
2813 GGAATGTTCTTGTGCTTCCATCAATGCTTAGCTTACCTTATGATTTGGATGCTTATGCT 2872  
2761 TGTGTGAACAGAACATGATTAATTTTGTCCAAATATTTACATTTTGGATTTATCTGACT 2820  
2873 TGTGTGAACAGAACATGATTAATTTTGTCCAAATATTTACATTTTGGATTTATCTGACT 2932  
2821 AAAAAAGAAACAACAGTACAGATTAATTTGAAAAAGAAAGCTCTTTCATTTATGAGCAAGT 2880  
2933 AAAAAAGAAACAACAGTACAGATTAATTTGAAAAAGAAAGCTCTTTCATTTATGAGCAAGT 2992  
2881 TACCAAGAGAACTTCTCTATCAGAGGAGAGAGTCTTTTCAAGTCTTTTGGAAATTAAT 2940  
2993 TACCAAGAGAACTTCTCTATCAGAGGAGAGAGTCTTTTCAAGTCTTTTGGAAATTAAT 3052  
2941 GACCTTCTGGAGACTTGTGTGTCAGCTTTTGTTTAAGATGTTTCCCTGAAGCCGAT 3000  
3053 GACCTTCTGGAGACTTGTGTGTCAGCTTTTGTTTAAGATGTTTCCCTGAAGCCGAT 3112  
3001 CCTTACATAGATTAATGATCAGAAATGTTTACACAGAACATACACTTGGCTTTAAAGCAAT 3060  
3113 CCTTACATAGATTAATGATCAGAAATGTTTACACAGAACATACACTTGGCTTTAAAGCAAT 3172  
3061 CAGAAATCAACGGTGAATTTTGGGATCAGAGAAAGTGAATTCATAGTGAAGCTTCAAGT 3120  
3173 CAGAAATCAACGGTGAATTTTGGGATCAGAGAAAGTGAATTCATAGTGAAGCTTCAAGT 3232  
3121 GGCATTAAGTCAAGTACACTTACAGCTATATTTGAATTTCTTCCCTGGGATTAAGA 3180  
3233 GGCATTAAGTCAAGTACACTTACAGCTATATTTGAATTTCTTCCCTGGGATTAAGA 3292  
3181 AAGTATCAGCTTAACATGATGTGCAAGAGTCTATCCATTTTGTGAGTCTGAATTCAGT 3240  
3293 AAGTATCAGCTTAACATGATGTGCAAGAGTCTATCCATTTTGTGAGTCTGAATTCAGT 3352  
3241 AGAGGAATTTGAGCAATTAATTAATCTTACCTTAACTTAATGCAATGTCTCAGTGGGG 3300

3353 AGAGGAATTTGAGCAATTAATTAATCTTACCTTAACTTAATGCAATGTCTCAGTGGGG 3412  
3301 AGTCTTAAGCCAGAGAGAGCTTTGAATATGCTACTGTGAGACAGAACAAAGAGTGGC 3360  
3413 AGTCTTAAGCCAGAGAGAGCTTTGAATATGCTACTGTGAGACAGAACAAAGAGTGGC 3472  
3361 ATGCAATTTGGGATGATCAGAGTCCAAACCTTTGACTCTTGGACCAAGCTCTCCCTG 3420  
3473 ATGCAATTTGGGATGATCAGAGTCCAAACCTTTGACTCTTGGACCAAGCTCTCCCTG 3532  
3421 GATATTAAGTCAAGCTTATGCACTGCTCTCACTCTTAACTTAATTTGAGCTTCTGAG 3480  
3533 GATATTAAGTCAAGCTTATGCACTGCTCTCACTCTTAACTTAATTTGAGCTTCTGAG 3592  
3481 GGAATCCCAATTAATGAGTGGCTTAAGCAGCAAGAAATAGCTTGGTGTGTTTGCATCT 3540  
3593 GGAATCCCAATTAATGAGTGGCTTAAGCAGCAAGAAATAGCTTGGTGTGTTTGCATCT 3652  
3541 ACTGAGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAAGCTTAAATGATCA 3600  
3653 ACTGAGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAAGCTTAAATGATCA 3712  
3601 GAAAGGCAAAATATCCAAAGTGAACGCTGACGGGGCTTACCTCACAAGTCTCTCTCT 3651  
3713 GAAAGGCAAAATATCCAAAGTGAACGCTGACGGGGCTTACCTCACAAGTCTCTCTCT 3772  
3652 -----CTTGGTGTGATCAGCCA 3669  
3773 CTGATTTGACACACAAACCGCTTATCTCTTACAGACAGACAGTCTGTGTGTGACCCA 3832  
3670 ATGACAGTTAAATATTTCCGCAAAATGTTTGAATTTGCTAATTTGTCAGCTCAATGTTGTA 3729  
3833 ATGACAGTTAAATATTTCCGCAAAATGTTTGAATTTGCTAATTTGTCAGCTCAATGTTGTA 3892  
3730 TATTAATGAAAGCTTCTGGCTCTTCTAGAAAGCAAGAAATCTTAAATTAAGAAAGCC 3789  
3893 TATTAATGAAAGCTTCTGGCTCTTCTAGAAAGCAAGAAATCTTAAATTAAGAAAGCC 3952  
3790 TTTGATTTAGATGTTGCTGTAAAGAAATTAAGATGATCTCATCATGATGATTTGAAT 3849  
3953 TTTGATTTAGATGTTGCTGTAAAGAAATTAAGATGATCTCATCATGATGATTTGAAT 4012  
3850 GTGTGTAACAAGCTTTTTCGGGCTCCGGGTAGAGTGGCATGCTCTTATGAAATTAACCTTA 3909  
4013 GTGTGTAACAAGCTTTTTCGGGCTCCGGGTAGAGTGGCATGCTCTTATGAAATTAACCTTA 4072  
3910 TTTAAGTGGCTTTATGAGTCTTCAAGAAAGCAATTTCTGTGACGAGACAGTGAAGAAAGTG 3969  
4073 TTTAAGTGGCTTTATGAGTCTTCAAGAAAGCAATTTCTGTGACGAGACAGTGAAGAAAGTG 4132  
3970 GAATATGATCATGGAAGAACTCAACCTTATTTAGATTTCTGAATGAAACCCAGTTTGT 4029  
4133 GAATATGATCATGGAAGAACTCAACCTTATTTAGATTTCTGAATGAAACCCAGTTTGT 4192  
4030 GTTAATATTTCTGCTGTGAGAAAATTTAAAGTTTCAAAATACCAAGATGCTTCAAGTCTC 4089  
4193 GTTAATATTTCTGCTGTGAGAAAATTTAAAGTTTCAAAATACCAAGATGCTTCAAGTCTC 4252  
4090 ATAGTGAATTAATTAAGCCAGAGACAGGCGGTGAGAAATTAATTAAGTGAAGTGAAG 4149  
4253 ATAGTGAATTAATTAAGCCAGAGACAGGCGGTGAGAAATTAATTAAGTGAAGTGAAG 4312  
4150 CTGTCTCTCTGTGACCTTTGACAGTATGTCAGAGGCTGCGCTCTTGTGAGATGAGCT 4209  
4313 CTGTCTCTCTGTGACCTTTGACAGTATGTCAGAGGCTGCGCTCTTGTGAGATGAGCT 4372  
4210 TCAAGCTCCCATCATCACTTCTCAGTCAATTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4269  
4373 TCAAGCTCCCATCATCACTTCTCAGTCAATTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4432  
4270 ATGGAATTTGGCTGTGA 4287  
4433 ATGGAATTTGGCTGTGA 4450

RESULT 8  
 AAL49815  
 ID AAL49815 standard; cDNA; 4335 BP.  
 XX  
 AC AAL49815;  
 XX  
 DT 04-DEC-2002 (first entry)  
 XX  
 DE Human platelet alloantigen Gova coding sequence.  
 XX  
 KM Human; platelet alloantigen; Gova; Govb; single nucleotide polymorphism;  
 KM SNP; biallelic; bleeding disorder; post-transfusion purpura;  
 KM post-transfusion platelet refractoriness; haemostatic; vaccine;  
 KM neonatal alloimmune thrombocytopenia; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..4335  
 FT /tag= a  
 FT /product= "Gova"  
 FT /partial  
 FT /note= "no stop codon"  
 FT replace(2108,C)  
 FT variation /tag= b  
 XX  
 PN MO200270738-A2.  
 XX  
 PD 12-SEP-2002.  
 XX  
 PF 07-MAR-2002; 2002WO-CA00291.  
 XX  
 PR 07-MAR-2001; 2001US-273941P.  
 XX  
 PA (SCHU/) SCHUH A.  
 XX  
 PI Schuh A, Ouwehand W;  
 XX  
 DR MPI; 2002-713460/77.  
 XX  
 DR P-PSDB; AAO19372.  
 XX  
 PT New isolated oligonucleotide binding to a region of CD109 nucleic acid  
 PT having a single nucleotide polymorphism that distinguishes a Gova  
 PT and/or Govb allele, useful for treating blood disorders e.g. alloimmune  
 PT thrombocytopenia  
 XX  
 PS Claim 6; Page 23-29; 69pp; English.  
 XX  
 CC The present invention relates to a sequence capable of binding  
 CC specifically to a CD109 nucleic acid which has a single nucleotide  
 CC polymorphism that distinguishes the Gova and Govb alleles. Detection of  
 CC the Gov genotype is useful for detecting whether the subject has or is at  
 CC risk of a blood disease, disorder or abnormal physical state, such as  
 CC bleeding, or increased risk of bleeding, due to alloimmune destruction of  
 CC blood platelets, e.g. post-transfusion purpura, post-transfusion  
 CC platelet refractoriness or neonatal alloimmune thrombocytopenia. The  
 CC nucleic acid and polypeptide are useful for Gov genotyping or phenotyping  
 CC individuals. The present sequence is the Gova coding sequence.  
 CC  
 SQ Sequence 4335 BP; 1296 A; 837 C; 922 G; 1280 T; 0 other;  
 Query Match 98.5%; Score 4221.4; DB 24; Length 4335;  
 Best Local Similarity 98.8%; Pred. No. 0;  
 Matches 4283; Conservative 0; Mismatches 1; Indels 51; Gaps 1;  
 Oy 1 ATGAGAGGCGCCAGCTCTGACCGCGCCACCTCTGCGGTGACGCGCGGCTG 60  
 Db 1 ATGAGAGGCGCCAGCTCTGACCGCGCCACCTCTCTGCGGTGACGCGCGGCTG 60  
 Oy 61 GCGGTGCTCCCGGCGCTCGGTTCTGTGACAGCCCGAGGATCATCAGGCCCGAGGA 120  
 Db 1141 ACTCTTGAAGAAAGAAATATGTAGTCATTAACGTGACACAGAAACTATATCTGAG 1200

Db 61 GCGGTGCTCCCGGCGCTCGGTTCTGTGACAGCCCGAGGATCATCAGGCCCGAGGA 120  
 Oy 121 AATGTGACTATTGGGGGTGAGCTTCTGGAAACACTGCCCTTCAAGGTGACTGTGAAGCG 180  
 Db 121 AATGTGACTATTGGGGGTGAGCTTCTGGAAACACTGCCCTTCAAGGTGACTGTGAAGCG 180  
 Oy 181 GAGCTGCTCAAGACAGCATCAAACTCTGCTCTGTCTGTGAAGCAGAAAGCTCTT 240  
 Db 181 GAGCTGCTCAAGACAGCATCAAACTCTGCTCTGTCTGTGAAGCAGAAAGCTCTT 240  
 Oy 241 GAAAAAGGCTCTTTTAAAGCACTTACTCTTCATCTACTCTTGAACAGTGCATGAG 300  
 Db 241 GAAAAAGGCTCTTTTAAAGCACTTACTCTTCATCTACTCTTGAACAGTGCATGAG 300  
 Oy 301 ATTATGAGCTACGTGTACCGGACGTACCGAGTAGATTTTATTCCTTAATGTATCC 360  
 Db 301 ATTATGAGCTACGTGTACCGGACGTACCGAGTAGATTTTATTCCTTAATGTATCC 360  
 Oy 361 CGCTTATCATTTAGACCAAGAGAAATATCTGTCTTCAATTCAAACAGAGGCTTATAC 420  
 Db 361 CGCTTATCATTTAGACCAAGAGAAATATCTGTCTTCAATTCAAACAGAGGCTTATAC 420  
 Oy 421 AAGCCAAAGCAAGAGTGAAGTTTGGCATTTGTACACTCTTCTCAGATTTTAAAGCTTAC 480  
 Db 421 AAGCCAAAGCAAGAGTGAAGTTTGGCATTTGTACACTCTTCTCAGATTTTAAAGCTTAC 480  
 Oy 481 AAACCTCTTTAAACATTTCTCATTAAGAGCCCAATTCATTAATTCACAGTGTG 540  
 Db 481 AAACCTCTTTAAACATTTCTCATTAAGAGCCCAATTCATTAATTCACAGTGTG 540  
 Oy 541 TCACAAACAAAGTATCTTGAAGTCAATTTCCAAACTTTTACAGTATCTTCCATCAATA 600  
 Db 541 TCACAAACAAAGTATCTTGAAGTCAATTTCCAAACTTTTACAGTATCTTCCATCAATA 600  
 Oy 601 CTGTGTGACTGTCTTATTCAGTTCAAGTGAATGACCAAGCATATTAATCATTTAG 660  
 Db 601 CTGTGTGACTGTCTTATTCAGTTCAAGTGAATGACCAAGCATATTAATCATTTAG 660  
 Oy 661 GTTTCAGAAATATGATTAACCAAAATTTGAAGGACTTTGACAGACCACTTAATGTTCT 720  
 Db 661 GTTTCAGAAATATGATTAACCAAAATTTGAAGGACTTTGACAGACCACTTAATGTTCT 720  
 Oy 721 ATGAATTCATGACATTTAATGTTACATCAAGCAAGTATATCATTTGGAGGACAGT 780  
 Db 721 ATGAATTCATGACATTTAATGTTACATCAAGCAAGTATATCATTTGGAGGACAGT 780  
 Oy 781 AAAGGAGAGTACGCTTACATTTTACCTTTATCTTTGGGGAAGAAAGAAATAT 840  
 Db 781 AAAGGAGAGTACGCTTACATTTTACCTTTATCTTTGGGGAAGAAAGAAATAT 840  
 Oy 841 ACAAACATTTAATGATTAATGATCTGCAAACTCTCTTTAATGATGAAGATGAA 900  
 Db 841 ACAAACATTTAATGATTAATGATCTGCAAACTCTCTTTAATGATGAAGATGAA 900  
 Oy 901 AATGTAATGATTTCTTCAATATGACTTTCTGAAATCTGATCTTCTCCCTGAGCA 960  
 Db 901 AATGTAATGATTTCTTCAATATGACTTTCTGAAATCTGATCTTCTCCCTGAGCA 960  
 Oy 961 GTAGAAATTTTAAACCAAGTACAGAAATCAGTTACAGGATTTTCAAGAAATGAAGACT 1020  
 Db 961 GTAGAAATTTTAAACCAAGTACAGAAATCAGTTACAGGATTTTCAAGAAATGAAGACT 1020  
 Oy 1021 AATGTGTTCTTCAACACATGATTAATCATTTGATTTTGAATTAATCACTGTCTG 1080  
 Db 1021 AATGTGTTCTTCAACACATGATTAATCATTTGATTTTGAATTAATCACTGTCTG 1080  
 Oy 1081 AAGCATCTTCAACTTCAAGCCACTGTGAAGTAACTGTGCTGATGAGCAACCACTG 1140  
 Db 1081 AAGCATCTTCAACTTCAAGCCACTGTGAAGTAACTGTGCTGATGAGCAACCACTG 1140  
 Oy 1141 ACTCTTGAAGAAAGAAATATGTAGTCATTAACGTGACACAGAAACTATATCTGAG 1200  
 Db 1141 ACTCTTGAAGAAAGAAATATGTAGTCATTAACGTGACACAGAAACTATATCTGAG 1200







```

Db 3361 ATGCATTTCTGGGTGTCATGACGATCAAACTTTTCTGACTCTCTGGAGCAGCTCCCTG 3420
QY 3421 GATATTTGAGTTGACAGCTATGCACTGCTCTCACTCTTCAATTTTCAAGCTTCTGAG 3480
Db 3421 GATATTTGAGTTGACAGCTATGCACTGCTCTCACTCTTCAATTTTCAAGCTTCTGAG 3480
QY 3481 GGAATCCCAATTAATGAGTGGCTTAAGCAGCAAGAAAGAAATAGCTTGGGTGTTTGATCT 3540
Db 3481 GGAATCCCAATTAATGAGTGGCTTAAGCAGCAAGAAATAGCTTGGGTGTTTGATCT 3540
QY 3541 ACTCAGGATACCACTGCTGCTTTAAAGGCTCTGCTGATTTTGGACCTTAATGATATCA 3600
Db 3541 ACTCAGGATACCACTGCTGCTTTAAAGGCTCTGCTGATTTTGGACCTTAATGATATCA 3600
QY 3601 GAAAGGACAATATCCAGTGAACCGGAGCGGGCTTACAGTCAACAAAGTCT----- 3651
Db 3601 GAAAGGACAATATCCAGTGAACCGGAGCGGGCTTACAGTCAACAAAGTCT----- 3651
QY 3652 -----CTTGCTGTGGTACAGCCA 3669
Db 3661 CTGATTGACACACAAACCGCTTACTCTTCAAGACAGAGCTTGTGTTGATACAGCCA 3720
QY 3670 ATGGCAGTTAATATTTCCGCAATGTTTGGATTTGCTATTTGTCAGCTCAATGTTGA 3729
Db 3721 ATGGCAGTTAATATTTCCGCAATGTTTGGATTTGCTATTTGTCAGCTCAATGTTGA 3780
QY 3730 TATTAATGAGAGGCTTGGGCTTCTAGAACGAGAGATCTATCCAAATCAAGAACCC 3789
Db 3781 TATTAATGAGAGGCTTGGGCTTCTAGAACGAGAGATCTATCCAAATCAAGAACCC 3840
QY 3790 TTTGATTTAGATGTTGCTGTAAGAAATTAAGATCTCATCATGATGATTTGAAT 3849
Db 3841 TTTGATTTAGATGTTGCTGTAAGAAATTAAGATCTCATCATGATGATTTGAAT 3900
QY 3850 GTGTGTACAGCTTTTCCGAGCCCGGAGTGAAGTGGAGTCTTTATGGAAGTTAACTTA 3909
Db 3901 GTGTGTACAGCTTTTCCGAGCCCGGAGTGAAGTGGAGTCTTTATGGAAGTTAACTTA 3960
QY 3910 TTTAAGTGGCTTATGCTGCTTCAAGAGCAATTTCTCTAGAGGAGCACTGAAAGAAAGT 3969
Db 3961 TTTAAGTGGCTTATGCTGCTTCAAGAGCAATTTCTCTAGAGGAGCACTGAAAGAAAGT 4020
QY 3970 GAATATGATCATGAGAAACTCAACCTCTATTTAGATTTCTGTAATGAACCCAGTTTGT 4029
Db 4021 GAATATGATCATGAGAAACTCAACCTCTATTTAGATTTCTGTAATGAACCCAGTTTGT 4080
QY 4030 GTTATATATTCCTGCTGTGAGAACTTTAAAGTTTCAATATCCCAAGATGCTTCAAGTCC 4089
Db 4081 GTTATATATTCCTGCTGTGAGAACTTTAAAGTTTCAATATCCCAAGATGCTTCAAGTCC 4140
QY 4090 ATATGATGATTTACTATATGAGCAAGAGAGACAGGCGGTGAGAAAGTTTACACTCTGAAGTGA 4149
Db 4141 ATATGATGATTTACTATATGAGCAAGAGAGACAGGCGGTGAGAAAGTTTACACTCTGAAGTGA 4200
QY 4150 CTGTCTCTCTGAGACCTTTGAGTGAATGCAAGGAGTGCAGCTTGTGAGAGATGAGAGT 4209
Db 4201 CTGTCTCTCTGAGACCTTTGAGTGAATGCAAGGAGTGCAGCTTGTGAGAGATGAGAGT 4260
QY 4210 TCAGGCTCCCATCATCACTCTTCACTATTTTATTTTCTGTTTCAAGCTTCTGATCTT 4269
Db 4261 TCAGGCTCCCATCATCACTCTTCACTATTTTATTTTCTGTTTCAAGCTTCTGATCTT 4320
QY 4270 ATGGAACCTTGGCTG 4284
Db 4321 ATGGAACCTTGGCTG 4335

```

RESULT 9  
 AAL49816  
 ID AAL49816 standard; cDNA; 4335 BP.  
 XX  
 AC AAL49816;

```

XX 04-DEC-2002 (first entry)
DT Human platelet alloantigen Govb coding sequence.
XX
DE Human, platelet allantoigen; Gova; Govb; single nucleotide polymorphism;
XX SNP; Ballelic; bleeding disorder; post-transfusion purpura;
XX post-transfusion platelet refractoriness; haemostatic; vaccine;
XX neonatal alloimmune thrombocytopenia; gene; ss.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..4335
FT FT /tag= a
FT FT /product= "Govb"
FT FT /partial
FT FT /note= "no stop codon"
FT FT replace(2108,A)
FT variation
FT FT /tag= b
XX
PN MO200270738-A2.
XX
XX 12-SEP-2002.
PD
XX
XX 07-MAR-2002; 2002MO-CA00291.
PF
XX
XX 07-MAR-2001; 2001US-273941P.
PR
XX
XX (SCHU/) SCHUH A.
XX
XX Schuh A, Ouwehand W;
XX PI
XX MPI. 2002-713460/77.
XX P-PSDB; AA019373.
XX
XX New isolated oligonucleotide binding to a region of CD109 nucleic acid
XX PT having a single nucleotide polymorphism that distinguishes a Gova
XX PT and/or Govb allele, useful for treating blood disorders e.g. alloimmune
XX PT thrombocytopenia
XX
XX Claim 7; Page 35-41; 69pp; English.
XX
XX The present invention relates to a sequence capable of binding
XX specifically to a CD109 nucleic acid which has a single nucleotide
XX polymorphism that distinguishes the Gova and Govb alleles. Detection of
XX the Govb genotype is useful for detecting whether the subject has or is at
XX risk of a blood disease, disorder or abnormal physical state, such as
XX bleeding, or increased risk of bleeding, due to alloimmune destruction of
XX blood platelets, e.g., post-transfusion purpura, post-transfusion
XX platelet refractoriness or neonatal alloimmune thrombocytopenia. The
XX CC nucleic acid and polypeptide are useful for Gov genotyping or phenotyping
XX CC individuals. The present sequence is the Govb coding sequence.
XX
SQ Sequence 4335 BP; 1295 A; 838 C; 922 G; 1280 T; 0 other:

```

```

Query Match 98.4%; Score 4219.8; DB 24; Length 4335;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 4282; Conservative 0; Mismatches 2; Indels 51; Gaps 1;

```

```

QY 1 ATGAGGGCCCAACCGCTCTGACCGCGCCCAACCTCTCTGAGTGAACCGCGGCTG 60
Db 1 ATGAGGGCCCAACCGCTCTGACCGCGCCCAACCTCTCTGAGTGAACCGCGGCTG 60
QY 61 GCCGTGCTCCCGGCGCTCTGAGTGAACCGCGCCCAACCTCTCTGAGTGAACCGCGGCTG 120
Db 61 GCCGTGCTCCCGGCGCTCTGAGTGAACCGCGCCCAACCTCTCTGAGTGAACCGCGGCTG 120
QY 121 AATGATCTATTTGGGAGTGAACCTCTGAGTGAACCGCGCCCAACCTCTCTGAGTGAACCGCG 180
Db 121 AATGATCTATTTGGGAGTGAACCTCTGAGTGAACCGCGCCCAACCTCTCTGAGTGAACCGCG 180
QY 181 GAGTGTCTCAAGACGATCAACCTCTCTCTGAGTGAACCGCGCCCAACCTCTCTGAGTGAACCGCG 240

```

181 GAGTGTCAAGACAGCATCAAACTCACTGTCTCTGCTGGAAGAGGAGTCTT 240  
241 GAAAAAGGCTCTTTAAAGACACTTACTTCCATCACTACCTGGAAGAGGAGATAG 300  
241 GAAAAAGGCTCTTTAAAGACACTTACTTCCATCACTACCTGGAAGAGGAGATAG 300  
301 ATTATGAGCTACGTGTAAACCGAGGATACCGAGATGAGATTTTATCTTAATAGTACC 360  
301 ATTATGAGCTACGTGTAAACCGAGGATACCGAGATGAGATTTTATCTTAATAGTACC 360  
361 CGCTTATATTTGAGACCAAGAGATATCTGTCTTCAATTCAGACAGAGAGGCTTATAC 420  
361 CGCTTATATTTGAGACCAAGAGATATCTGTCTTCAATTCAGACAGAGAGGCTTATAC 420  
421 AAGCCAAAGAGAGAGTGAAGTTGCGCATGTGTACACTCTCTGAGATTTTAAAGCTTAC 480  
421 AAGCCAAAGAGAGAGTGAAGTTGCGCATGTGTACACTCTCTGAGATTTTAAAGCTTAC 480  
481 AAAACCTCTTTAAACATTTCTCATTAAGGACCCCAATCAATTTGATCCACAGTGTG 540  
481 AAAACCTCTTTAAACATTTCTCATTAAGGACCCCAATCAATTTGATCCACAGTGTG 540  
541 TCACAAACAAAGTATCTTGAGTCAATTCGAAAATTTTCAAGCTATCTTCCATCCATA 600  
541 TCACAAACAAAGTATCTTGAGTCAATTCGAAAATTTTCAAGCTATCTTCCATCCATA 600  
601 CTGGTGACGTGCTATCAAGTTCAAGTGAATGACGACATTTTATCAATCAATTCAG 660  
601 CTGGTGACGTGCTATCAAGTTCAAGTGAATGACGACATTTTATCAATCAATTCAG 660  
661 GTTTCAGAAATATGATTAACCAAAATTTGAGTGAATTTGACAGACCACTTATTTGTTCT 720  
661 GTTTCAGAAATATGATTAACCAAAATTTGAGTGAATTTGACAGACCACTTATTTGTTCT 720  
721 ATGATTTTAAGCATTTAATGATGATCACTACGCGAAAGTATACATATGGAAGCAGTG 780  
721 ATGATTTTAAGCATTTAATGATGATCACTACGCGAAAGTATACATATGGAAGCAGTG 780  
781 AAGAGAGCGTAAAGCTTACATTTTACCTTATCTTTGAGGAAAGAAATATTT 840  
781 AAGAGAGCGTAAAGCTTACATTTTACCTTATCTTTGAGGAAAGAAATATTT 840  
841 ACAAACATTTAAGATTAATGATGATGCAAACTTCTTTTAAATGATGAAGATGAAA 900  
841 ACAAACATTTAAGATTAATGATGATGCAAACTTCTTTTAAATGATGAAGATGAAA 900  
901 AATGTAATGATTTCTTCAATGAGACTTTCTGAATACCTGATCTATCTTCCCTGAGCA 960  
901 AATGTAATGATTTCTTCAATGAGACTTTCTGAATACCTGATCTATCTTCCCTGAGCA 960  
961 GTAGAAATTTTAAACACAGTGAAGATCAAGTATCAAGTATTTTAAAGAAATGTAAGCACT 1020  
961 GTAGAAATTTTAAACACAGTGAAGATCAAGTATCAAGTATTTTAAAGAAATGTAAGCACT 1020  
1021 AATGTTCTTCAAGCAATGATTAATCATGATGATTTTGAATTAATCACTAGTCTTG 1080  
1021 AATGTTCTTCAAGCAATGATTAATCATGATGATTTTGAATTAATCACTAGTCTTG 1080  
1081 AAGCATCTCTCAACTTCAAGCACTGTAAGGATCTGCTGATGAGCAACCACTG 1140  
1081 AAGCATCTCTCAACTTCAAGCACTGTAAGGATCTGCTGATGAGCAACCACTG 1140  
1141 ACTCTGAAAGAAAGAAATTAATGATGATCAATCAAGTGAACAGAGAAATTAATAGTAC 1200  
1141 ACTCTGAAAGAAAGAAATTAATGATGATCAATCAAGTGAACAGAGAAATTAATAGTAC 1200  
1201 TACTGAGCGGATCTAAACAGTGAAGATCAAGAAATGAGAGCTGTTCAAGAAATTAAT 1260  
1201 TACTGAGCGGATCTAAACAGTGAAGATCAAGAAATGAGAGCTGTTCAAGAAATTAAT 1260  
1261 ACTGTCCCAAGTGAAGCTTTTAAATGATTTCCCAATCTGAGAGATTCAGTGAAG 1320  
1261 ACTGTCCCAAGTGAAGCTTTTAAATGATTTCCCAATCTGAGAGATTCAGTGAAG 1320

1261 ACTGTCCCAAGTGAAGCTTTTAAATGATTTCCCAATCTGAGAGATTCAGTGAAG 1320  
1321 CTACAGTTGAAGGCTATTTCTTGGTAAAGTAAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1380  
1321 CTACAGTTGAAGGCTATTTCTTGGTAAAGTAAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1380  
1381 AAGTCTCTTAAGATATCAATCACTTAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1440  
1381 AAGTCTCTTAAGATATCAATCACTTAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1440  
1441 TCGCTTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1500  
1441 TCGCTTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1500  
1501 GTATCCAGGAGGACAGTGTGCTGTAAGGAAACAAATTTCAACAAATTTCTTAAACA 1560  
1501 GTATCCAGGAGGACAGTGTGCTGTAAGGAAACAAATTTCAACAAATTTCTTAAACA 1560  
1561 CCAAGAAATTTCTGAGCTCCAAAGCTGTGTAATGTGATTAATGAAGTGAAGTGAAG 1620  
1561 CCAAGAAATTTCTGAGCTCCAAAGCTGTGTAATGTGATTAATGAAGTGAAGTGAAGTGAAG 1620  
1621 GAAATTAATGATGATGTTCTTAAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680  
1621 GAAATTAATGATGATGTTCTTAAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680  
1681 CTATATTTGAAGTAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1740  
1681 CTATATTTGAAGTAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1740  
1741 ACAAGCTGACTCCATAGTGGAGTGTAGCTGTTGACAAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1800  
1741 ACAAGCTGACTCCATAGTGGAGTGTAGCTGTTGACAAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1800  
1801 GCTCTTAATGATTTCAATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1860  
1801 GCTCTTAATGATTTCAATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1860  
1861 TATTTATTTAGGATGATTTCAATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1920  
1861 TATTTATTTAGGATGATTTCAATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1920  
1921 TTGACAGATGCAAACTCAGAAAGATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1980  
1921 TTGACAGATGCAAACTCAGAAAGATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1980  
1981 GCTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2040  
1981 GCTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2040  
2041 GTAGAGTCCATGATGCGAAAGCAATTTCCAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2100  
2041 GTAGAGTCCATGATGCGAAAGCAATTTCCAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2100  
2101 ATGGGTTTACAGATTTTACCAAGATTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2160  
2101 ATGGGTTTACAGATTTTACCAAGATTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2160  
2161 GTGGCTTACTGTTTGTGATCTGAGAGACTGAGGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2220  
2161 GTGGCTTACTGTTTGTGATCTGAGAGACTGAGGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2220  
2221 GAGCTCAAGCTTCCAAACATTTTCAATTTTGAATCTTCTGACTCTGTTACAG 2280  
2221 GAGCTCAAGCTTCCAAACATTTTCAATTTTGAATCTTCTGACTCTGTTACAG 2280  
2281 GGTGAAGATTTGCTTGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2340  
2281 GGTGAAGATTTGCTTGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2340  
2341 AAGTATATCTTAAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2400  
2341 AAGTATATCTTAAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2400

QY 2401 GCCACAGGCCACAGACACCTTCTGGTCCAGTAGAGATGAGGCAACTGTTCTTTT 2460  
 Db 2401 GCCACAGGCCACAGACACCTTCTGGTCCAGTAGAGATGAGGCAACTGTTCTTTT 2460  
 QY 2461 CCCATAGGCGCAACATCTGGAGAAATTCCTATCAAGTCAAGCTCTTCAACCCT 2520  
 Db 2461 CCCATAGGCGCAACATCTGGAGAAATTCCTATCAAGTCAAGCTCTTCAACCCT 2520  
 QY 2521 GCTTCTGATGCTGCAACCGATGATTTTGTAAAGCTGAGAAATGAAATATCATAT 2580  
 Db 2521 GCTTCTGATGCTGCAACCGATGATTTTGTAAAGCTGAGAAATGAAATATCATAT 2580  
 QY 2581 TCACAAATTCATCTTATTAGACTTGAAGTCAATAGGCTACAGATACCTGAAACCTTG 2640  
 Db 2581 TCACAAATTCATCTTATTAGACTTGAAGTCAATAGGCTACAGATACCTGAAACCTTG 2640  
 QY 2641 AGTTTCTCATTTCTCTCTATACAGTGAAGTGAAGAAAGATTCAATCACTGCAAT 2700  
 Db 2641 AGTTTCTCATTTCTCTCTATACAGTGAAGTGAAGAAAGATTCAATCACTGCAAT 2700  
 QY 2701 GGAGATGTTCTTGGTCTTCCATCAATGGCTTAGCCCATGATGGATGGCTTATGGC 2760  
 Db 2701 GGAGATGTTCTTGGTCTTCCATCAATGGCTTAGCCCATGATGGATGGCTTATGGC 2760  
 QY 2761 TGTGTGAACAGAACATGATTAATTTTGTCTCAAAATTTTATCACTTATCTGACT 2820  
 Db 2761 TGTGTGAACAGAACATGATTAATTTTGTCTCAAAATTTTATCACTTATCTGACT 2820  
 QY 2821 AAAAAGAAACATGACAGATTAATTTGAAAGAAAGAAAGCTTTTCAATTTATGAGCAAGT 2880  
 Db 2821 AAAAAGAAACATGACAGATTAATTTGAAAGAAAGAAAGCTTTTCAATTTATGAGCAAGT 2880  
 QY 2881 TACACAGAGAACTTCTATCAGAGGAGAAAGCTTTCAGTCTTTTGGAAATTTAT 2940  
 Db 2881 TACACAGAGAACTTCTATCAGAGGAGAAAGCTTTCAGTCTTTTGGAAATTTAT 2940  
 QY 2941 GACCTTCTGGAGCACTTGTGTGCTGAGCTTTTGTAAAGATGTTTCTTGAAGCCGAT 3000  
 Db 2941 GACCTTCTGGAGCACTTGTGTGCTGAGCTTTTGTAAAGATGTTTCTTGAAGCCGAT 3000  
 QY 3001 CCTTACATAGATTTGATCAGATGATGTTTACAGAAACATACCTTGGCTTAAAGACAT 3060  
 Db 3001 CCTTACATAGATTTGATCAGATGATGTTTACAGAAACATACCTTGGCTTAAAGACAT 3060  
 QY 3061 CAGAAATCCAAAGCTGAAATTTTGGATCCAGAAAGATGATTAATGATGAGCTTCAAGT 3120  
 Db 3061 CAGAAATCCAAAGCTGAAATTTTGGATCCAGAAAGATGATTAATGATGAGCTTCAAGT 3120  
 QY 3121 GGCATTAAGATGCTGATACCTTACAGCCCTATATTTGTAATCTTCCCTGGGATTTAG 3180  
 Db 3121 GGCATTAAGATGCTGATACCTTACAGCCCTATATTTGTAATCTTCCCTGGGATTTAG 3180  
 QY 3181 AAGTATCAGCTTAACATGATGCTGAGGCTTATTAATTTTGGAGTCTGAATCAGT 3240  
 Db 3181 AAGTATCAGCTTAACATGATGCTGAGGCTTATTAATTTTGGAGTCTGAATCAGT 3240  
 QY 3241 AAGGGAATTTCAACAATTTACTTACAGCCCTTATTAATTTTGGAGTCTGAATCAGT 3300  
 Db 3241 AAGGGAATTTCAACAATTTACTTACAGCCCTTATTAATTTTGGAGTCTGAATCAGT 3300  
 QY 3301 AGTCTTAAGCGAAGGAGCTTGAATATGCTGAGCTTGAAGCAGAAACAAGAAAGTGGC 3360  
 Db 3301 AGTCTTAAGCGAAGGAGCTTGAATATGCTGAGCTTGAAGCAGAAACAAGAAAGTGGC 3360  
 QY 3361 ATGCAATTTCTGGTGTATCAGAGTCCAAATCTTCTGACTCTCGGAGCCACGCTCCCTG 3420  
 Db 3361 ATGCAATTTCTGGTGTATCAGAGTCCAAATCTTCTGACTCTCGGAGCCACGCTCCCTG 3420  
 QY 3421 GATATTTGAAGTTCAGGCTTATGACAGTCTTCACTTTTCAATTTTCAAGCTTCTGAG 3480  
 Db 3421 GATATTTGAAGTTCAGGCTTATGACAGTCTTCACTTTTCAATTTTCAAGCTTCTGAG 3480

QY 3481 GGAATCCCAATATATGAGTGGCTTAAGCAGGCAAGAAATAGCTGGTGGTTTGCATCT 3540  
 Db 3481 GGAATCCCAATATATGAGTGGCTTAAGCAGGCAAGAAATAGCTGGTGGTTTGCATCT 3540  
 QY 3541 ACTCAGATATCAGCTGTGCTTTAAAGGCTCTGTCTGAATTTTGGAGCCCTTAATGATACA 3600  
 Db 3541 ACTCAGATATCAGCTGTGCTTTAAAGGCTCTGTCTGAATTTTGGAGCCCTTAATGATACA 3600  
 QY 3601 GAAAGGACAAATATCCAAAGTGAACGCTGACGAGGCTAGCTCACAAAGTCTT----- 3651  
 Db 3601 GAAAGGACAAATATCCAAAGTGAACGCTGACGAGGCTAGCTCACAAAGTCTTAAAGTTT 3660  
 QY 3652 -----CTTCTGTGTACAGCCA 3669  
 Db 3661 CTGATTTGACACACACACACCGCTTACTCTTCCAGACAGAGCTTGTGTGTAACGCCA 3720  
 QY 3670 ATGGCAGTTAATATTTCCGCAATGCTTTTGGATTTGCTATTTTGCAGTCAATGTTGTA 3729  
 Db 3721 ATGGCAGTTAATATTTCCGCAATGCTTTTGGATTTGCTATTTTGCAGTCAATGTTGTA 3780  
 QY 3730 TTAATATGAAAGGCTTGGGCTTCTGAAAGAGAAATATCAATCAATCAAGAGCC 3789  
 Db 3781 TTAATATGAAAGGCTTGGGCTTCTGAAAGAGAAATATCAATCAATCAAGAGCC 3840  
 QY 3790 TTTGATTTAGATGTTGCTGTAAGAAATAAAGATGATCTCAATCATGAGATTTGAAT 3849  
 Db 3841 TTTGATTTAGATGTTGCTGTAAGAAATAAAGATGATCTCAATCATGAGATTTGAAT 3900  
 QY 3850 GTGTGTACAAGCTTTTGGGCTTGGAGTGAAGTGCATGCTCTTATGAAATTAACCTA 3909  
 Db 3901 GTGTGTACAAGCTTTTGGGCTTGGAGTGAAGTGCATGCTCTTATGAAATTAACCTA 3960  
 QY 3910 TTAAGTGGCTTATGAGTGTCTTCAAGACAAATTTCTGAGGAGACAGTGAAGAAAGT 3969  
 Db 3961 TTAAGTGGCTTATGAGTGTCTTCAAGACAAATTTCTGAGGAGACAGTGAAGAAAGT 4020  
 QY 3970 GAATATGATCATGAAACTCAACTTATTAGATTTCTGTAATGAATGAACCAAGTTTGT 4029  
 Db 4021 GAATATGATCATGAAACTCAACTTATTAGATTTCTGTAATGAATGAACCAAGTTTGT 4080  
 QY 4030 GTTAATATTTCTGCTGTGAGAAACTTTAAAGTTTCAATCAATCAAGATGCTTCAAGTTC 4089  
 Db 4081 GTTAATATTTCTGCTGTGAGAAACTTTAAAGTTTCAATCAATCAAGATGCTTCAAGTTC 4140  
 QY 4090 ATAGTGAATTAATAGACCAAGAGACAGAGGCTGAGAAATTAACCTTGAAGTGAAG 4149  
 Db 4141 ATAGTGAATTAATAGACCAAGAGACAGAGGCTGAGAAATTAACCTTGAAGTGAAG 4200  
 QY 4150 CTGCTCTCTGTGACCTTTGAGATGATCCAGGCTGCGCTTGTGAGGATGAGAT 4209  
 Db 4201 CTGCTCTCTGTGACCTTTGAGATGATCCAGGCTGCGCTTGTGAGGATGAGAT 4260  
 QY 4210 TCAGGCTCCCATCATCACTCTTCAAGTCAATTTTCTGTTTCAAGCTTCTGACTTT 4269  
 Db 4261 TCAGGCTCCCATCATCACTCTTCAAGTCAATTTTCTGTTTCAAGCTTCTGACTTT 4320  
 QY 4270 ATGGAATTTGGCTG 4284  
 Db 4321 ATGGAATTTGGCTG 4335

## RESULT 10

ID AAA62010  
 ID AAA62010 standard; DNA; 4473 BP.

AC AAA62010;

DT 02-FEB-2001 (first entry)

DE Hydrophobic domain protein cDNA HP02837 isolated from HT-1080 cells.  
 XX Human; secreted protein; membrane protein; hydrophobic domain;  
 KW proliferation control; differentiation induction; material transport;

KM biophylaxis; signal receptor; ion channel; transporter; immunostimulant;  
 KM immunosuppressant; haematopoiesis regulator; chemotactic; chemokinetic;  
 KM haemostatic; thrombolytic; anti-inflammatory; tumour inhibition;  
 KM autoimmune disease; Alzheimer's disease; Parkinson's disease; cancer;  
 KM gene therapy; ss.  
 OS Homo sapiens.  
 PN WO200029448-A2.  
 PD 25-MAY-2000.  
 XX 17-NOV-1999; 99WO-JP06412.  
 XX 17-NOV-1998; 98JP-0326255.  
 XX 22-DEC-1998; 98JP-0364315.  
 XX 16-MAR-1999; 99JP-0069811.  
 XX 27-APR-1999; 99JP-0119299.  
 XX 19-MAY-1999; 99JP-01138169.  
 XX (SAGA ) SAGAMI CHEM RES CENT.  
 PA (PROT-) PROTEGENE INC.  
 XX Kato S, Kimura T;  
 PI WPI; 2000-387753/33.  
 DR P-PSDB; AAB12127.  
 XX Proteins comprising hydrophobic regions, such as secretory and membrane  
 PT proteins, useful in research and diagnostics and having various  
 PT activities e.g. immunomodulatory, anti-inflammatory, chemokinetic,  
 PT hemostatic, thrombolytic -  
 PS Claim 4; Page 253-261; 410pp; English.  
 XX Secretory proteins play important roles in the proliferation control, the  
 CC differentiation induction, the material transport and the biophylaxis of  
 CC cells. Membrane proteins have important roles as signal receptors, ion  
 CC channels and transporters. The present sequence is the coding sequence  
 CC for a human protein which has at least one hydrophobic domain. The  
 CC protein encoded by the present sequence may be a secretory or a membrane  
 CC protein. The encoded protein may have cytokine and cell  
 CC proliferation/differentiation activity, immune stimulating or suppressing  
 CC activity, haematopoiesis activity, tissue growth activity,  
 CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, anti-inflammatory activity and tumour  
 CC inhibition activity. The present sequence could therefore be used for  
 CC treatment of autoimmune disease, Alzheimer's disease, Parkinson's  
 CC disease, and cancer via gene therapy.  
 CC  
 XX Sequence 4473 BP; 1328 A; 867 C; 952 G; 1326 T; 0 other;  
 SQ  
 Query Match 98.4%; Score 4219.6; DB 21; Length 4473;  
 Bsect Local Similarity 98.7%; Freq. No. 0;  
 Matches 4283; Conservative 0; Mismatches 4; Indels 51; Gaps 1;

DB 285 GAAAAAGGCTCTTTAAGACACTTACTCTTCATCTCTGAAACAGTGAATGAG 344  
 QY 301 ATTATGAGCTACGTGTAACCGGACGTACAGAGATGAGATTTATCTTAATAGTACC 360  
 DB 345 ATTTAGAGCTAGCTGTAACCGGACGTACAGAGATGAGATTTATCTTAATAGTACC 404  
 QY 361 CGCTTATCATTTTGAACCAAGAGATATCTGCTTCAATCAACAGCAAGGCTTATAC 420  
 DB 405 CGCTTATCATTTTGAACCAAGAGATATCTGCTTCAATCAACAGCAAGGCTTATAC 464  
 QY 421 AAGCCAAAGCAAGATGAACTTGGATTTGATACCTGCTGCAATTTTAAAGCTTAC 480  
 DB 465 AAGCCAAAGCAAGATGAACTTGGATTTGATACCTGCTGCAATTTTAAAGCTTAC 524  
 QY 481 AAAACCTCTTTAAACATCTCTATTAAGAACCCCAATCAAAATTTGATCCAAAGTGTG 540  
 DB 525 AAAACCTCTTTAAACATCTCTATTAAGAACCCCAATCAAAATTTGATCCAAAGTGTG 584  
 QY 541 TCACAACAAAGATCTTGGAGTCATTTCCAAACTTTTCAAGTATCTTCCATCCAAAT 600  
 DB 585 TCACAACAAAGATCTTGGAGTCATTTCCAAACTTTTCAAGTATCTTCCATCCAAAT 644  
 QY 601 CTGGTGACCTGCTATTCAGATTCAGATGAATGACCAACATATATCATATTTTACG 660  
 DB 645 CTGGTGACCTGCTATTCAGATTCAGATGAATGACCAACATATATCATATTTTACG 704  
 QY 661 GTTTCGAATATGATATTAACCAAAATTTGAAGTACTTTGAGACACCATATATGTTCT 720  
 DB 705 GTTTCGAATATGATATTAACCAAAATTTGAAGTACTTTGAGACACCATATATGTTCT 764  
 QY 721 ATGAATTTCAAGCATTTTAAATGATGATACATACGCGCAAGATATCATATGGAAGCCAGTG 780  
 DB 765 ATGAATTTCAAGCATTTTAAATGATGATACATACGCGCAAGATATCATATGGAAGCCAGTG 824  
 QY 781 AAGGAGAGCTAACGCTTACATTTTACCTTTATCTTTGGGAAAGAAATATAT 840  
 DB 825 AAGGAGAGCTAACGCTTACATTTTACCTTTATCTTTGGGAAAGAAATATAT 884  
 QY 841 ACAAACATTTTAAATGATGATGATGCAAACTTCTTTTAAATGATGATGATGAA 900  
 DB 885 ACAAACATTTTAAATGATGATGATGCAAACTTCTTTTAAATGATGATGATGAA 944  
 QY 901 AATGATATGATTTCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
 DB 945 AATGATATGATTTCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1004  
 QY 961 GTTGAATTTTAAACCAAGTACAGATGATGATGATGATGATGATGATGATGATGATGAT 1020  
 DB 1005 GTTGAATTTTAAACCAAGTACAGATGATGATGATGATGATGATGATGATGATGATGAT 1064  
 QY 1021 AATGATGATTTCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080  
 DB 1065 AATGATGATTTCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1124  
 QY 1081 AAGCCATCTCTCAACTTCAAGCAAGCACTGTAAGTAACTGCTGCTGATGCAACCAACTG 1140  
 DB 1125 AAGCCATCTCTCAACTTCAAGCAAGCACTGTAAGTAACTGCTGCTGATGCAACCAACTG 1184  
 QY 1141 ACTCTTGAAGAAAGAAATTAATGATGATTAAGTATTAAGTATTAAGTATTAAGTATTAAG 1200  
 DB 1185 ACTCTTGAAGAAAGAAATTAATGATGATTAAGTATTAAGTATTAAGTATTAAGTATTAAG 1244  
 QY 1201 TACTGAGGAGATCTTAACAGTGAATTCAGAAATGGAAGCTGTTAGAAATTAATTAAT 1260  
 DB 1245 TACTGAGGAGATCTTAACAGTGAATTCAGAAATGGAAGCTGTTAGAAATTAATTAAT 1304  
 QY 1261 ACTGCTCCCAAGTGAAGCTTTTAAATGATTAATCCATCTGAGAGATTCAGATGAG 1320  
 DB 1305 ACTGCTCCCAAGTGAAGCTTTTAAATGATTAATCCATCTGAGAGATTCAGATGAG 1364  
 QY 1321 CTACAGTTGAAGGCTTATTTCTTGGTATGAAAGTATGATGATGATGATGATGATGATGATGAT 1380

Db 1365 CTACAGTTGAAGCCCTATTTCTTGGTAGTAAAAAGTAGACATGGCACTTCATAGTCTGTTT 1424  
Qy 1381 AAGTCTCTAGTAAAGACATACATCCAACTAAAAACAAGAGATGAATAATATAAGGTGGGA 1440  
Db 1425 AAGTCTCTAGTAAAGACATACATCCAACTAAAAACAAGAGATGAATAATATAAGGTGGGA 1484  
Qy 1441 TGGCTTTTGAAGTGGTGGTAAAGTGGCAACAAAGATGAAGAGTTAAGCTAATATGTA 1500  
Db 1485 TGGCTTTTGAAGTGGTGGTAAAGTGGCAACAAAGATGAAGAGTTAAGCTAATATGTA 1544  
Qy 1501 GATATCAGGGGACAGTTGGTGGTGTAGAGAAAACAAATTCACAAATGTTCTTTAAACA 1560  
Db 1545 GATATCAGGGGACAGTTGGTGGTGTAGAGAAAACAAATTCACAAATGTTCTTTAAACA 1604  
Qy 1561 CCAGAAAATTTCTTGAATCCAAAGGCTGTGTAATTGTGATATATATGAAGATGATGGG 1620  
Db 1605 CCAGAAAATTTCTTGAATCCAAAGGCTGTGTAATTGTGATATATATGAAGATGATGGG 1664  
Qy 1621 GAAATTTAAGATGATGTTCTAAATAATTCCTGTTCAAGCTGTTTTTAAAAATAGATMAAG 1680  
Db 1665 GAAATTTAAGATGATGTTCTAAATAATTCCTGTTCAAGCTGTTTTTAAAAATAGATMAAG 1724  
Qy 1681 CTATATGAGATAAGTGAAGTGAACCATCTGAGAAAGTCTCTTAGATCTGTG 1740  
Db 1725 CTATATGAGATAAGTGAAGTGAACCATCTGAGAAAGTCTCTTAGATCTGTG 1784  
Qy 1741 ACACAGCTGATCTCAATGATGGGATTTGATGCTTGACAAAAGTGAATCTGATGAAT 1800  
Db 1785 ACACAGCTGATCTCAATGATGGGATTTGATGCTTGACAAAAGTGAATCTGATGAAT 1844  
Qy 1801 GCCTCAATGATATTAACAATGAAAAATGTGTCCATGAGTTGGAATTTATTAACACAGA 1860  
Db 1845 GCCTCAATGATATTAACAATGAAAAATGTGTCCATGAGTTGGAATTTATTAACACAGA 1904  
Qy 1861 TATTTATTTAGGCATGTTTCATGAATCTTTTGCAAGTCTTTACAGAAATGTGACTGTGGTA 1920  
Db 1905 TATTTATTTAGGCATGTTTCATGAATCTTTTGCAAGTCTTTACAGAAATGTGACTGTGGTA 1964  
Qy 1921 TTGACAGATGCAAACTCCGAAAGATTTATGATGGTGTATGACAAATGACAGAAAT 1980  
Db 1965 TTGACAGATGCAAACTCCGAAAGATTTATGATGGTGTATGACAAATGACAGAAAT 2024  
Qy 1981 GCTGAGAGTTTATGAGAGAAAATGAAGACATATTTAGATATTCATGACTTTCTTGG 2040  
Db 2025 GCTGAGAGTTTATGAGAGAAAATGAAGACATATTTAGATATTCATGACTTTCTTGG 2084  
Qy 2041 GGTAGAGTCCACATGTCGAAAGCAATTTCCAGACACTTGAATTTGGCTAGACACCAAC 2100  
Db 2085 GGTAGAGTCCACATGTCGAAAGCAATTTCCAGACACTTGAATTTGGCTAGACACCAAC 2144  
Qy 2101 ATGGGTTACAGATTTTACCAAGAAATTTGAAGTACGTACCTGATTTCTATCATTCTTGG 2160  
Db 2145 ATGGGTTACAGATTTTACCAAGAAATTTGAAGTACGTACCTGATTTCTATCATTCTTGG 2204  
Qy 2161 GTGGCTACTGTTTGTGATCTCTGAGACCTGGGCTTGGACATAACACTACTCCAGTG 2220  
Db 2205 GTGGCTACTGTTTGTGATCTCTGAGACCTGGGCTTGGACATAACACTACTCCAGTG 2264  
Qy 2221 GAGCTCCAAAGCCTTCCAAACATTTTCAATTTTTTGAATCTTCCACTCTGTTATCAGA 2280  
Db 2265 GAGCTCCAAAGCCTTCCAAACATTTTCAATTTTTTGAATCTTCCACTCTGTTATCAGA 2324  
Qy 2281 GGTGAAGATTTTCTTGGAAAAATTAATATTTCAATTTATTTGAAAAAGTCCACTGAGGT 2340  
Db 2325 GGTGAAGATTTTCTTGGAAAAATTAATATTTCAATTTATTTGAAAAAGTCCACTGAGGT 2384  
Qy 2341 AAGGTATCATTTGAGAAAAAGTGAACAATTTGATATTTCTAATGACTTCAAGTGAATTAAT 2400  
Db 2385 AAGGTATCATTTGAGAAAAAGTGAACAATTTGATATTTCTAATGACTTCAAGTGAATTAAT 2444  
Qy 2401 GCCACAGGCCACACAGACACCTTGTGTTCCAGTGAAGATGGGCAACTGTTCTTTT 2460  
Db 2445 GCCACAGGCCACACAGACACCTTGTGTTCCAGTGAAGATGGGCAACTGTTCTTTT 2504

Qy 2461 CCATCAGGCCAACACATCTGGAGAAATTCCTATCAGTCAAGCTCTTTCACCCACT 2520  
Db 2505 CCATCAGGCCAACACATCTGGAGAAATTCCTATCAGTCAAGCTCTTTCACCCACT 2564  
Qy 2521 GCTTCTGATGCTGTCAACCAGATGATTTTAAAGGCTGAAGGAATGAATAATCATAT 2580  
Db 2565 GCTTCTGATGCTATCACCCAGATGATTTTAAAGGCTGAAGGAATGAATAATCATAT 2624  
Qy 2581 TCACATTCATCTTATTTAGACTTGAACATAATAGCTACAGAGTACCCCTGAAAACTTGG 2640  
Db 2625 TCACATTCATCTTATTTAGACTTGAACATAATAGCTACAGAGTACCCCTGAAAACTTGG 2684  
Qy 2641 AGTTCTCATTTCCCTCAATACAGTGACTGAGAGTGAAGAAAGTTCAATCAGTCAAT 2700  
Db 2685 AGTTCTCATTTCCCTCAATACAGTGACTGAGAGTGAAGAAAGTTCAATCAGTCAAT 2744  
Qy 2701 GAGATGTTCTTGGTCTTTCATCAATGAGCTTAACTTGAATTCGATGCTTATGGC 2760  
Db 2745 GAGATGTTCTTGGTCTTTCATCAATGAGCTTAACTTGAATTCGATGCTTATGGC 2804  
Qy 2761 TGTGTGAACAGACATGATTAATTTTGTCTCAAAATTTTACATTTTGGATTTCTGACT 2820  
Db 2805 TGTGTGAACAGACATGATTAATTTTGTCTCAAAATTTTACATTTTGGATTTCTGACT 2864  
Qy 2821 AAAAGAAACAACTGACAGATTAATTTGAAGAAAGGCTTTTCAATTTATGAGGCAAGT 2880  
Db 2865 AAAAGAAACAACTGACAGATTAATTTGAAGAAAGGCTTTTCAATTTATGAGGCAAGT 2924  
Qy 2881 TACAGAGAACTTCTCTATCAGAGGAAAGTGGCTTTCAAGTCTTTTGGAAATTA 2940  
Db 2925 TACAGAGAACTTCTCTATCAGAGGAAAGTGGCTTTCAAGTCTTTTGGAAATTA 2984  
Qy 2941 GACCTTCTGGAGACCTTGTGTGACCTTTTGTAAAGTGTTCCTTGAAGCCGAT 3000  
Db 2985 GACCTTCTGGAGACCTTGTGTGACCTTTTGTAAAGTGTTCCTTGAAGCCGAT 3044  
Qy 3001 CCTTACATGATTTGATCAGAAATGTGTACAGAAATCACTTGGCTTTAAAGGAAT 3060  
Db 3045 CCTTACATGATTTGATCAGAAATGTGTGTACAGAAATCACTTGGCTTTAAAGGAAT 3104  
Qy 3061 CAGAAATCCAAAGGTAATTTTGGATCAGAGAGAGATTCATAGAGCTTCAAGT 3120  
Db 3105 CAGAAATCCAAAGGTAATTTTGGATCAGAGAGAGATTCATAGAGCTTCAAGT 3164  
Qy 3121 GGCATTAAGGTCAGTAACCTTACAGCTTATATTTGTAATCTTCTCGGGAATATAGA 3180  
Db 3165 GGCATTAAGGTCAGTAACCTTACAGCTTATATTTGTAATCTTCTCGGGAATATAGA 3224  
Qy 3181 AAGTATCAGCCTAACATTTGATGAGTGAAGAGTCTATCCATTTTGGAGTCTGAATTCAGT 3240  
Db 3225 AAGTATCAGCCTAACATTTGATGAGTGAAGAGTCTATCCATTTTGGAGTCTGAATTCAGT 3284  
Qy 3241 AGAGGAATTTCAACAATTTATCTAGCCCTTATTAATTTGATGACTGTCACTAGTGGGG 3300  
Db 3285 AGAGGAATTTCAACAATTTATCTAGCCCTTATTAATTTGATGACTGTCACTAGTGGGG 3344  
Qy 3301 AGTCTTAAGCGAAGAAAGCTTTGAATATGCTGACTTGAAGACGAAACAAGAGTGGC 3360  
Db 3345 AGTCTTAAGCGAAGAAAGCTTTGAATATGCTGACTTGAAGACGAAACAAGAGTGGC 3404  
Qy 3361 ATGCAATTTCTGGGTCTCATGAGAGTCCAACTTTTGTGACTCTGGCAGCCAGCTCCCTG 3420  
Db 3405 ATGCAATTTCTGGGTCTCATGAGAGTCCAACTTTTGTGACTCTGGCAGCCAGCTCCCTG 3464  
Qy 3421 GATATTTGAAGTTGACGCTTATGACTGCTCACAATTTTCAATTTTCAAGCTTCTGAG 3480  
Db 3465 GATATTTGAAGTTGACGCTTATGACTGCTCACAATTTTCAAGCTTCTTCAAGCTTCTGAG 3524  
Qy 3481 GGAATCCAAATTTATGAGTGGCTTAAGCAGGCAAGAAATAGCTTGGTGTGTTTGCATCT 3540  
Db 3525 GGAATCCAAATTTATGAGTGGCTTAAGCAGGCAAGAAATAGCTTGGTGTGTTTGCATCT 3584

QY 3541 ACTGAGATACCACTGGGCTTTTAAAGGCTCTGTCTGAATTTGACGCCCTTAATGATACA 3600  
 DB 3585 ACTGAGATACCACTGGGCTTTTAAAGGCTCTGTCTGAATTTGACGCCCTTAATGATACA 3644  
 QY 3601 GAAAGGAAATATCCAGTGAACCTGACGGGGCTTACTACCAAGTCTT----- 3651  
 DB 3645 GAAAGGAAATATCCAGTGAACCTGACGGGGCTTACTACCAAGTCTTAAAGTTT 3704  
 QY 3652 -----CTTGTCTGTGTACAGCA 3669  
 DB 3705 CTGATGACACACACACCGCTTACTCTTGCAGACAGACAGCTTGTCTGTGTACAGCA 3764  
 QY 3670 ATGCGATTAATATTTCCGCAATATGTTTGGATTGCTTAATTTGTCACTCAATATGTTGA 3729  
 DB 3765 ACGGAGATTAATTTCCGCAATATGTTTGGATTGCTTAATTTGTCACTCAATATGTTGA 3824  
 QY 3730 TATATGTGAAGGCTTCTGGGCTTCTTGAAGACGAAATCTTATCCAAATCAAGAAAGCC 3789  
 DB 3825 TATATGTGAAGGCTTCTGGGCTTCTTGAAGACGAAATCTTATCCAAATCAAGAAAGCC 3884  
 QY 3790 TTTGATTTAGATGTTGCTGTAAAGAAATATAAGATGATCTCAATCATGTGATTTGAAT 3849  
 DB 3885 TTTGATTTAGATGTTGCTGTAAAGAAATATAAGATGATCTCAATCATGTGATTTGAAT 3944  
 QY 3850 GTGTGTACAACTTTTGGGCCCCGGGTAGAGGTGCATGCTCTTATGAAATTAACTTA 3909  
 DB 3945 GTGTGTACAACTTTTGGGCCCCGGGTAGAGGTGCATGCTCTTATGAAATTAACTTA 4004  
 QY 3910 TTAAGTGGCTTTATGTGTCTTCAAGACATTTCTGTGACGAGACAGTGAAGAAAGTG 3969  
 DB 4005 TTAAGTGGCTTTATGTGTCTTCAAGACATTTCTGTGACGAGACAGTGAAGAAAGTG 4064  
 QY 3970 GAATATGATCANGAAACTCAACCTTATTTAGATCTGTAAATGAAACCAAGTTTGT 4029  
 DB 4065 GAATATGATCANGAAACTCAACCTTATTTAGATCTGTAAATGAAACCAAGTTTGT 4124  
 QY 4030 GTTAATATCTCTGTGTGAGAACTTTAAAGTTTCAATCCAAAGATCTTCAAGTCTC 4089  
 DB 4125 GTTAATATCTCTGTGTGAGAACTTTAAAGTTTCAATCCAAAGATCTTCAAGTCTC 4184  
 QY 4080 ATAGTGAATTAATGAGCCCAAGACAGAGCGGTGAGAAATTAATCTGAAAGTGAAG 4149  
 DB 4185 ATAGTGAATTAATGAGCCCAAGACAGAGCGGTGAGAAATTAATCTGAAAGTGAAG 4244  
 QY 4150 CTGTCTCTCTGTGACCTTTGAGATGATGCCAGGCTGCCGCTTGTGAGATGAGAGCT 4209  
 DB 4245 CTGTCTCTCTGTGACCTTTGAGATGATGCCAGGCTGCCGCTTGTGAGATGAGAGCT 4304  
 QY 4210 TCAGGCTCCCATCATCACTCTTCAAGTCAATTTTATTTCTGTTCAAGCTTCTGACTTT 4269  
 DB 4305 TCAGGCTCCCATCATCACTCTTCAAGTCAATTTTATTTCTGTTCAAGCTTCTGACTTT 4364  
 QY 4270 ATGGAATTTGGCTGTGA 4287  
 DB 4365 ATGGAATTTGGCTGTGA 4382

## RESULT 11

AAA60199 standard; DNA; 4335 BP.

AAA60199;

02-FEB-2001 (first entry)

Hydrophobic domain protein cDNA HP02837 isolated from HT-1080 cells.

Human; secreted protein; membrane protein; hydrophobic domain;  
 proliferation control; differentiation induction; material transport;  
 biophysics; signal receptor; ion channel; transporter; immunostimulant;  
 immunosuppressant; haematopoiesis regulator; chemokine; chemokine;  
 haemostatic; thrombolytic; anti-inflammatory; tumour inhibition;  
 autoimmune disease; Alzheimer's disease; Parkinson's disease; cancer;

KM gene therapy; ss.  
 XX Homo sapiens.  
 OS Homo sapiens.  
 XX W0200029448-A2.  
 PD 25-MAY-2000.  
 XX 17-NOV-1999; 99MO-JP06412.  
 XX 17-NOV-1998; 98JP-0326255.  
 PR 22-DEC-1998; 98JP-0364315.  
 PR 16-MAR-1999; 99JP-0069811.  
 PR 27-APR-1999; 99JP-0119299.  
 PR 19-MAY-1999; 99JP-0138169.  
 XX (SAGA) SAGAMI CHEM RES CENT.  
 PA (PROT-) PROTEGENE INC.  
 XX Kato S, Kimura T;  
 PI WPI: 2000-387753/33.  
 DR P-PSDB; AAB12127.  
 XX Proteins comprising hydrophobic regions, such as secretory and membrane  
 PT proteins, useful in research and diagnostics and having various  
 PT activities e.g. immunomodulatory, antiinflammatory, chemokine,  
 PT hemostatic, thrombolytic -  
 XX Claim 3; Page 244-246; 410pp; English.  
 PS Secretory proteins play important roles in the proliferation control, the  
 CC differentiation induction, the material transport and the biophysics of  
 CC cells. Membrane proteins have important roles as signal receptors, ion  
 CC channels and transporters. The present sequence is the coding sequence  
 CC for a human protein which has at least one hydrophobic domain. The  
 CC protein encoded by the present sequence may be a secretory or a membrane  
 CC protein. The encoded protein may have cytokine and cell  
 CC proliferation/differentiation activity, immune stimulating or suppressing  
 CC activity, haematopoiesis activity, tissue growth activity,  
 CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, anti-inflammatory activity and tumour  
 CC inhibition activity. The present sequence could therefore be used for  
 CC treatment of autoimmune disease, Alzheimer's disease, Parkinson's  
 CC disease, and cancer via gene therapy.  
 XX Sequence 4335 BP; 1295 A; 838 C; 922 G; 1280 T; 0 other;  
 SQ  
 Query Match 98.4%; Score 4216.6; DB 21; Length 4335;  
 Best Local Similarity 98.7%; Pred. No. 0;  
 Matches 4280; Conservative 0; Mismatches 4; Indels 51; Gaps 1;  
 QY 1 ATGCAAGGGCCACCGCTCTGACCGCGCCACCTCTCTGTGTGACACCGCGGCTG 60  
 DB 1 ATGCAAGGGCCACCGCTCTGACCGCGCGCCACCTCTCTGTGTGACACCGCGGCTG 60  
 QY 61 GCCGTGGCTCCGCGGCTCGGTTTGTGTGACAGCCGAGGATCATCAGGCGGAGGA 120  
 DB 61 GCCGTGGCTCCGCGGCTCGGTTTGTGTGACAGCCGAGGATCATCAGGCGGAGGA 120  
 QY 121 AATGTACTATTTGGGGTGAAGCTTGTGAACACTGCGCTTCAAGGTGACTGGAAGCG 180  
 DB 121 AATGTACTATTTGGGGTGAAGCTTGTGAACACTGCGCTTCAAGGTGACTGGAAGCG 180  
 QY 181 GAGCTGTCAAGACAGATCAAACTCACTGTCTGTCTGTGAGACAGAGAGTCTTT 240  
 DB 181 GAGCTGTCAAGACAGATCAAACTCACTGTCTGTCTGTGAGACAGAGAGTCTTT 240  
 QY 241 GAAAGAGGCTTTTAAAGACATTAATCTTCACTCACTCACTGAAAGTGAAGTGAAG 300  
 DB 241 GAAAGAGGCTTTTAAAGACATTAATCTTCACTCACTCACTGAAAGTGAAGTGAAG 300  
 QY 301 ATTATGAGCTACGTGAACCGGACGTACCAGATGAGATTTATTTCTTAATGATACC 360

```

Db      301  ||||| 360
AtTTAGAGCTAGGTGTAACCGGACGTACCAGAGATGATTTATCTTAATAGTACC
Qy      361  ||||| 420
CGCTTATCTTTGAGACCAAGAAATATCTGCTTCAATCAACAGACAGGCTTAAC
Db      361  ||||| 420
CGCTTATCTTTGAGACCAAGAAATATCTGCTTCAATCAACAGACAGGCTTAAC
Qy      421  ||||| 480
AAGCCAAAGCAAGAAAGTGAAGTTGCGATGTTACACTCTTCCAGATTTTAAGCCTTAC
Db      421  ||||| 480
AAGCCAAAGCAAGAAAGTGAAGTTGCGATGTTACACTCTTCCAGATTTTAAGCCTTAC
Qy      481  ||||| 540
AAAACTCTTTAAACATCTCATTTAAGAGACCCCAATCAAAATTTGATCCAGATGTTG
Db      481  ||||| 540
AAAACTCTTTAAACATCTCATTTAAGAGACCCCAATCAAAATTTGATCCAGATGTTG
Qy      541  ||||| 600
TCACACAAAGTATCTTGAGTCACTTCCAAAACCTTTCAGCTATCTTCCATCCAAAT
Db      541  ||||| 600
TCACACAAAGTATCTTGAGTCACTTCCAAAACCTTTCAGCTATCTTCCATCCAAAT
Qy      601  ||||| 660
CTTGAGCTGCTGTATTCAGTTCAAGTGAATGACACAGATATTCATCATTTTCAG
Db      601  ||||| 660
CTTGAGCTGCTGTGTATTCAGTTCAAGTGAATGACACAGATATTCATCATTTTCAG
Qy      661  ||||| 720
GTTTCAGATATGATTTACCAAAATTTGAAGTGAAGTTCGACACACCATTAATGTTCT
Db      661  ||||| 720
GTTTCAGATATGATTTACCAAAATTTGAAGTGAAGTTCGACACACCATTAATGTTCT
Qy      721  ||||| 780
ATGAATCTTAAGATTTAATAGTGTACATCAAGGCAAAATATACATATGGAAGCCAGG
Db      721  ||||| 780
ATGAATCTTAAGATTTAATAGTGTACATCAAGGCAAAATATACATATGGAAGCCAGG
Qy      781  ||||| 840
AAAGAGACGTAAACGCTTACATTTTACCTTATCTTTTGGGGAAGAAGAAATAT
Db      781  ||||| 840
AAAGAGAGATTAACGCTTACATTTTACCTTATCTTTTGGGGAAGAAGAAATAT
Qy      841  ||||| 900
ACAAAAACATTTAAGTAATAGATCTGCAAACTTCTCTTTAATGATGAAGATGAAA
Db      841  ||||| 900
ACAAAAACATTTAAGTAATAGATCTGCAAACTTCTCTTTAATGATGAAGATGAAA
Qy      901  ||||| 960
AATGTAATGATTTCTTCAATGAGACTTCTGAAATACCTGATCTATCTTCCCTGACCA
Db      901  ||||| 960
AATGTAATGATTTCTTCAATGAGACTTCTGAAATACCTGATCTATCTTCCCTGACCA
Qy      961  ||||| 1020
GTAGAAATTTTAAACACAGTGAAGATGATTAAGATTTTCAAGAAATGTAAGCACT
Db      961  ||||| 1020
GTAGAAATTTTAAACACAGTGAAGATGATTAAGATTTTCAAGAAATGTAAGCACT
Qy      1021  ||||| 1080
AATGTTCTTCAAGACATGATTAATCATTTGAGTTTGTATTAATCTACTGCTTG
Db      1021  ||||| 1080
AATGTTCTTCAAGACATGATTAATCATTTGAGTTTGTATTAATCTACTGCTTG
Qy      1081  ||||| 1140
AAGCATCTCTCAATCTCAAGCACTGAGAAAGTAACTGCTGATGAGCAACCACTG
Db      1081  ||||| 1140
AAGCATCTCTCAATCTCAAGCACTGAGAAAGTAACTGCTGATGAGCAACCACTG
Qy      1141  ||||| 1200
ACTCTTGAAGAAAGAAATATATGATCATTAACAGTGAACAGAGAACTATTAAGT
Db      1141  ||||| 1200
ACTCTTGAAGAAAGAAATATATGATCATTAACAGTGAACAGAGAACTATTAAGT
Qy      1201  ||||| 1260
TACTGAGGCGATCTTAACAGTGAATCAGAAAATGGAAGCTGTTCAAGAAATTAAT
Db      1201  ||||| 1260
TACTGAGGCGATCTTAACAGTGAATCAGAAAATGGAAGCTGTTCAAGAAATTAAT
Qy      1261  ||||| 1320
ACTGTCCTCCCAAGTGAACCTTTAAGATTAATCCCAATCTGAGAGATTCAGTGA
Db      1261  ||||| 1320
ACTGTCCTCCCAAGTGAACCTTTAAGATTAATCCCAATCTGAGAGATTCAGTGA
Qy      1321  ||||| 1380
CTACAGTTGAAGGCTATTTCTTGATGAAGAAAGATGAGATGATTAATGATGCTGTT
Db      1321  ||||| 1380
CTACAGTTGAAGGCTATTTCTTGATGAAGAAAGATGAGATGATTAATGATGCTGTT
Qy      1381  ||||| 1440
AAGTCTCTAGTAAGCATACATCAATTAACAAAGATGAATTAATTAAGTGGGA
Db      1381  ||||| 1440
AAGTCTCTAGTAAGCATACATCAATTAACAAAGATGAATTAATTAAGTGGGA
Db      1441  ||||| 1500
TCGCTTTTGAAGTTGGTGTAGTGCAACAAAGATGAAGAGTTAAGTATAGTGA
Qy      1441  ||||| 1500
TCGCTTTTGAAGTTGGTGTAGTGCAACAAAGATGAAGAGTTAAGTATAGTGA
Db      1501  ||||| 1560
GTATCAGAGGACAGTTGGTGTAGTGCAACAAATTCACACATGTTCTTTTAAACA
Qy      1501  ||||| 1560
GTATCAGAGGACAGTTGGTGTAGTGCAACAAATTCACACATGTTCTTTTAAACA
Db      1561  ||||| 1620
CGAAGAAATTTTGGACTCCAAAAGCTGTGTATTTGATTAATTAAGAGATGATGG
Qy      1561  ||||| 1620
CGAAGAAATTTTGGACTCCAAAAGCTGTGTATTTGATTAATTAAGAGATGATGG
Db      1621  ||||| 1680
GAATTAATAGTATGTTCTTAATAATTCCTGTACGTTGTTTAAATTAAGTAAAG
Qy      1621  ||||| 1680
GAATTAATAGTATGTTCTTAATAATTCCTGTACGTTGTTTAAATTAAGTAAAG
Db      1681  ||||| 1740
CTATATGAGATTAAGTGAAGCTGAACCATCTGAGAAAGTCTCTTGAAGATCTGTG
Qy      1681  ||||| 1740
CTATATGAGATTAAGTGAAGCTGAACCATCTGAGAAAGTCTCTTGAAGATCTGTG
Db      1741  ||||| 1800
ACACAGCTGACTCATAGTTGGAGTTGATGCTGTGAACAAAGTGAATCTGATGAT
Qy      1741  ||||| 1800
ACACAGCTGACTCATAGTTGGAGTTGATGCTGTGAACAAAGTGAATCTGATGAT
Db      1801  ||||| 1860
GCTCTAATGATTAATTAAGTGAAGTGTGCTGCAAGTGTGAACCTTATTAACAGGA
Qy      1801  ||||| 1860
GCTCTAATGATTAATTAAGTGAAGTGTGCTGCAAGTGTGAACCTTATTAACAGGA
Db      1861  ||||| 1920
TATTAATGAGCATGTTGATTAATCTTTCAGCTTTCAGAAATGAGATCTGAGTGA
Qy      1861  ||||| 1920
TATTAATGAGCATGTTGATTAATCTTTCAGCTTTCAGAAATGAGATCTGAGTGA
Db      1921  ||||| 1980
TTGACAGATGCAAACTTCAGAGAGATTAATGATGCTGTTATGACATGACAGATAT
Qy      1921  ||||| 1980
TTGACAGATGCAAACTTCAGAGAGATTAATGATGCTGTTATGACATGACAGATAT
Db      1981  ||||| 2040
GCTGAGAGTTTATGAGAGAAATGAAGAGATTAATGATGATTAATGACATGATCTT
Qy      1981  ||||| 2040
GCTGAGAGTTTATGAGAGAAATGAAGAGATTAATGATGATTAATGACATGATCTT
Db      2041  ||||| 2100
GATACAGTCAACATGTCGAGAAAGATTTTCAAGAGCTTGGATTTGCTGACACCAAC
Qy      2041  ||||| 2100
GATACAGTCAACATGTCGAGAAAGATTTTCAAGAGCTTGGATTTGCTGACACCAAC
Db      2101  ||||| 2160
ATGAGTTTCAAGATTTTACCAAGAAATTTGAAGTATGATGATCTATCACTTCTGG
Qy      2101  ||||| 2160
ATGAGTTTCAAGATTTTACCAAGAAATTTGAAGTATGATGATCTATCACTTCTGG
Db      2161  ||||| 2220
GTGCTACTGTTTGTATCTGAGAGACCTGGGCTTGGACATTAACACTCCAGTGG
Qy      2161  ||||| 2220
GTGCTACTGTTTGTATCTGAGAGACCTGGGCTTGGACATTAACACTCCAGTGG
Db      2221  ||||| 2280
GAGCTCCAAAGCTTCCAAACATTTTCTTGTGAATCTTCCCTACTCTGTTATGAGA
Qy      2221  ||||| 2280
GAGCTCCAAAGCTTCCAAACATTTTCTTGTGAATCTTCCCTACTCTGTTATGAGA
Db      2281  ||||| 2340
GGTGAAGATTTGCTTGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT
Qy      2281  ||||| 2340
GGTGAAGATTTGCTTGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT
Db      2341  ||||| 2400
AAGGTATCATTTGAGAAAGTGAACAAATTTATATTTCAATGACTTCAAGTGAATTAAT
Qy      2341  ||||| 2400
AAGGTATCATTTGAGAAAGTGAACAAATTTATATTTCAATGACTTCAAGTGAATTAAT
Db      2401  ||||| 2460
GCCACAGGCGACAGAGACCTTCTGGTCCAGTGAAGATGAGGCGAACCTGTTCTTTT
Qy      2401  ||||| 2460
GCCACAGGCGACAGAGACCTTCTGGTCCAGTGAAGATGAGGCGAACCTGTTCTTTT
Db      2461  ||||| 2520
CCCATCAGGCGACACATCTGGAGAAATTTCTATCAACAGTCAAGCTTTTCAACCACT
Qy      2461  ||||| 2520
CCCATCAGGCGACACATCTGGAGAAATTTCTATCAACAGTCAAGCTTTTCAACCACT
Db      2520  ||||| 2580
CCCATCAGGCGACACATCTGGAGAAATTTCTATCAACAGTCAAGCTTTTCAACCACT
Qy      2520  ||||| 2580
CCCATCAGGCGACACATCTGGAGAAATTTCTATCAACAGTCAAGCTTTTCAACCACT

```



QY 2521 GCTTCTGATGCTGTCACCCAGATGATTTTAAAGGCTGAAGAAATAGAAAAATCATAT 2580  
 Db 2521 GCTTCTGATGCTATCACCCAGATGATTTTAAAGGCTGAAGAAATAGAAAAATCATAT 2580  
 QY 2581 TCACATCATCTTATTATGACTTGACTGACAAATAGGCTCAAGAGTACCTGAAAACTTTG 2640  
 Db 2581 TCACATCATCTTATTATGACTTGACTGACAAATAGGCTCAAGAGTACCTGAAAACTTTG 2640  
 QY 2641 AGTTCTCATTTCCCTTAATACGTGACGTGAGTGAAGAGTACAGATACAGCAATT 2700  
 Db 2641 AGTTCTCATTTCCCTTAATACGTGACGTGAGTGAAGAGTACAGATACAGCAATT 2700  
 QY 2701 GGAAGTGTCTTGGCTCTTCATCAATGCTTAGCCTTATGATTCGATTCGATTCGCTTATG 2760  
 Db 2701 GGAAGTGTCTTGGCTCTTCATCAATGCTTAGCCTTATGATTCGATTCGATTCGCTTATG 2760  
 QY 2761 TGTGCTGAACAGAACATGATTAATTTGCTCCAAATATTACATTTTGATTAATCTGACT 2820  
 Db 2761 TGTGCTGAACAGAACATGATTAATTTGCTCCAAATATTACATTTTGATTAATCTGACT 2820  
 QY 2821 AAAAAGAAACAACGACAGATTAATTTGAAGAAAGCTTTGATTTATGAGGCAAGT 2880  
 Db 2821 AAAAAGAAACAACGACAGATTAATTTGAAGAAAGCTTTGATTTATGAGGCAAGT 2880  
 QY 2881 TACAGAGAGAACTTCTCTATCAGAGGAGAGATGCTCTTCACTGCTTTTGGAAATTAAT 2940  
 Db 2881 TACAGAGAGAACTTCTCTATCAGAGGAGAGATGCTCTTCACTGCTTTTGGAAATTAAT 2940  
 QY 2941 GACCTTCTGGAGACCTTGGTGTCACTTTTGTATTAAGATGTTTCTTTGAAGCCGAT 3000  
 Db 2941 GACCTTCTGGAGACCTTGGTGTCACTTTTGTATTAAGATGTTTCTTTGAAGCCGAT 3000  
 QY 3001 CCTTACATAGATTTGATCAGATGCTGTACAGAAACATACATGCTGCTTAAAGGACAT 3060  
 Db 3001 CCTTACATAGATTTGATCAGATGCTGTACAGAAACATACATGCTGCTTAAAGGACAT 3060  
 QY 3061 CAGAAATCCAAAGGTGATTTTGGGATCCAGAAAGATGATTCATTAAGTCAAGT 3120  
 Db 3061 CAGAAATCCAAAGGTGATTTTGGGATCCAGAAAGATGATTCATTAAGTCAAGT 3120  
 QY 3121 GGCATTAATAAGCTCAGTAACTTAAAGCTTAATTTGATCTTCTCTGGGATTAAGA 3180  
 Db 3121 GGCATTAATAAGCTCAGTAACTTAAAGCTTAATTTGATCTTCTCTGGGATTAAGA 3180  
 QY 3181 AAGTATCAGCTTACATTTGATGTCAGAGAGTCAATTTTGGAGTCTGAATTCAGT 3240  
 Db 3181 AAGTATCAGCTTACATTTGATGTCAGAGAGTCAATTTTGGAGTCTGAATTCAGT 3240  
 QY 3241 AGAGAAATTTTCAACAATTTACTCTAGCCCTTAATCTTATGATTCATCTAGTGGG 3300  
 Db 3241 AGAGAAATTTTCAACAATTTACTCTAGCCCTTAATCTTATGATTCATCTAGTGGG 3300  
 QY 3301 AGTCTTAAAGGAGAAAGCTTTTGAATATGCTGAATTTTGAAGTCTGAGTCTGCTG 3360  
 Db 3301 AGTCTTAAAGGAGAAAGCTTTTGAATATGCTGAATTTTGAAGTCTGAGTCTGCTG 3360  
 QY 3361 ATGCAATTTTGGGTGTATGAGATCCAAATTTTCTGATCTTGGAGCCAGCTCCCTG 3420  
 Db 3361 ATGCAATTTTGGGTGTATGAGATCCAAATTTTCTGATCTTGGAGCCAGCTCCCTG 3420  
 QY 3421 GATATTTGAAGTTGACGCTTATGCACTGTCTCAACAATTTTCAAAATTTTCAACTTCTG 3480  
 Db 3421 GATATTTGAAGTTGACGCTTATGCACTGTCTCAACAATTTTCAAAATTTTCAACTTCTG 3480  
 QY 3481 GGAATCCCAATTTATGAGTGTCTAAGCAGGCAAGAAATAGCTTGGTGTGTTTGCATCT 3540  
 Db 3481 GGAATCCCAATTTATGAGTGTCTAAGCAGGCAAGAAATAGCTTGGTGTGTTTGCATCT 3540  
 QY 3541 ACTCAGATACCACTGTGCTTTAAAGCTGTGCTGAATTTGACGCTTATTAATTAACA 3600  
 Db 3541 ACTCAGATACCACTGTGCTTTAAAGCTGTGCTGAATTTGACGCTTATTAATTAACA 3600

QY 3601 GAAAGACAAATATCCAGTACGCTGACGGGCTTAGCTCACCAAGTCTT----- 3651  
 Db 3601 GAAAGACAAATATCCAGTACGCTGACGGGCTTAGCTCACCAAGTCTT----- 3651  
 QY 3652 -----CTTGGCTGTGACAGCCA 3669  
 Db 3661 CTGATTTGACACACACACCGCTTACTCTTACAGACAGAGGCTTCTGTGTACAGCCA 3720  
 QY 3670 ATGCAATTAATTAATTTCCGCAATGTTTTGATTTGATTTGATTTGATTTGATTTGATTT 3729  
 Db 3721 ACGGCATTATATTTTCCGCAATGTTTTGATTTGATTTGATTTGATTTGATTTGATTT 3780  
 QY 3730 TATTAATGTAAGGCTTCTGAGCTTCTAGAGCGAAGATCTATCCAAATCAAGAGCC 3789  
 Db 3781 TATTAATGTAAGGCTTCTGAGCTTCTAGAGCGAAGATCTATCCAAATCAAGAGCC 3840  
 QY 3790 TTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 3849  
 Db 3841 TTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 3900  
 QY 3850 GTGTGTACAAAGCTTTTCCGGCCCGGTTAGAGTGCATGAGCTTTATGGAATTAACCTA 3909  
 Db 3901 GTGTGTACAAAGCTTTTCCGGCCCGGTTAGAGTGCATGAGCTTTATGGAATTAACCTA 3960  
 QY 3910 TTAAGTGTCTTTATGAGTGTCTTCAAGACAAATTTCTGAGCGAGACAGTGAAGAAAGT 3969  
 Db 3961 TTAAGTGTCTTTATGAGTGTCTTCAAGACAAATTTCTGAGCGAGACAGTGAAGAAAGT 4020  
 QY 3970 GAATATGATCATGGAAGAACTCAACCTCTATTTAGATTTCTGTAAATGAATCCAGTTTGT 4029  
 Db 4021 GAATATGATCATGGAAGAACTCAACCTCTATTTAGATTTCTGTAAATGAATCCAGTTTGT 4080  
 QY 4030 GTTAATATTCCTGCTGTGAGAACTTTAAAGTTTCAATATCCAAAGATGCTTCACTGCTC 4089  
 Db 4081 GTTAATATTCCTGCTGTGAGAACTTTAAAGTTTCAATATCCAAAGATGCTTCACTGCTC 4140  
 QY 4090 ATAGTGAATTAATAGAGCCAGAGACAGCGGTGAGAAATTAACAATCTGAAAGTGAAG 4149  
 Db 4141 ATAGTGAATTAATAGAGCCAGAGACAGCGGTGAGAAATTAACAATCTGAAAGTGAAG 4200  
 QY 4150 CTGTCTCTCTGTGACCTTTGCAATGATGTCAGAGGCTGCGCTTGTGAGATGAGAGCT 4209  
 Db 4201 CTGTCTCTCTGTGACCTTTGCAATGATGTCAGAGGCTGCGCTTGTGAGATGAGAGCT 4260  
 QY 4210 TCAGGCTCCCATGATGATCTTCAATGATTTTATTTTCAAGCTTCTGATCTT 4269  
 Db 4261 TCAGGCTCCCATGATGATCTTCAATGATTTTATTTTCAAGCTTCTGATCTT 4320  
 QY 4270 ATGGAATTTGGCTG 4284  
 Db 4321 ATGGAATTTGGCTG 4335

RESULT 12  
 AAD49436  
 ID AAD49436 standard; DNA; 4146 BP.  
 XX AAD49436;  
 AC AAD49436;  
 DT 24-MAR-2003 (first entry)  
 DE Human blood cell surface antigen, CD109 encoding DNA #1.  
 KW Human; GPI-anchored TGF-beta1 binding protein; r150 protein; cancer;  
 KW glycosylphosphatidylinositol; transforming growth factor-beta1;  
 KW therapy; blood cell surface antigen; CD109; gene; ds.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..4146  
 FT /tag= a  
 FT /product= "Human CD109 protein"

FT /transl\_except= (pos:2044..2046, aa:Xaa)  
FT /note= "Xaa corresponds to Ser, Tyr; No start  
FT and stop codon"  
FT [partial  
XX MO200285942-A2.  
XX 31-OCT-2002.  
XX 24-APR-2002; 2002MO-CA00560.  
XX 24-APR-2001; 2001US-285713P.  
XX 14-FEB-2002; 2002US-356163P.  
XX (UWMC-) UNIV MCGILL.  
XX Philip A, Tam B;  
XX WPI; 2003-093100/08.  
XX P-PSDB; AAE32014.  
XX Novel transforming growth factor (TGF)-beta 1 binding reagent which  
XX comprises r150 protein which acts as accessory receptor of TGF-beta,  
XX useful for negatively modulating TGF-beta activity, and thus for  
XX treating cancer -  
XX Claim 7; Page 109-112; 127pp; English.  
XX The invention relates to novel transforming growth factor (TGF)-beta 1  
XX binding reagent which comprises a glycosylphosphatidyl inositol (GPI)-  
XX anchored TGF-beta binding protein referred to as r150 which acts as  
XX an accessory receptor of TGF-beta. The invention is used for negatively  
XX modulating TGF-beta activity, and thus for treating conditions  
XX characterised by overproduction of TGF-beta, such as cancer. Antisense  
XX molecules of the invention are used for increasing TGF-beta availability  
XX and increase graft success. The present sequence is human blood cell  
XX surface antigen, CD109 encoding DNA. CD109 is a variant of r150 sequence.  
XX Sequence 4146 BP; 1260 A; 776 C; 885 G; 1224 T; 1 other;  
SQ  
Query Match 96.6%; Score 4140.2; DB 25; Length 4146;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 4142; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 64 GTGGCTCCCGGGCTCGGTTCTGGTGACAGCCCGAGGATCATGAGCCCGAGAGAAAT 123  
DB 1 GTGGCTCCCGGGCTCGGTTCTGGTGACAGCCCGAGGATCATGAGCCCGAGAGAAAT 60  
QY 124 GTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCAAGGTGATGTGAAGCGGAG 183  
DB 61 GTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCAAGGTGATGTGAAGCGGAG 120  
QY 184 CTGCTCAAGACAGCATCAAACTCACTGCTCTGCTCTGGAAGCAAGAGAGCTTTTGA 243  
DB 121 CTGCTCAAGACAGCATCAAACTCACTGCTCTGCTCTGGAAGCAAGAGAGCTTTTGA 180  
QY 244 AAAAGCTCTTTAAGACCTTACTCTTCATCACTCACTCAAGAGTGAAGTGAAGT 303  
DB 181 AAAAGCTCTTTAAGACCTTACTCTTCATCACTCACTCAAGAGTGAAGTGAAGT 240  
QY 304 TATGAGCTACGTGTAAACCGAGTACCCAGATGAGATTTATTTCTCTATATGATCCCGC 363  
DB 241 TATGAGCTACGTGTAAACCGAGTACCCAGATGAGATTTATTTCTCTATATGATCCCGC 300  
QY 364 TTATCATTTTGAAGACCAAGAAATATCTGCTCTTCAATTCAAACAGACAGAGCTTTATCAAG 423  
DB 301 TTATCATTTTGAAGACCAAGAAATATCTGCTCTTCAATTCAAACAGACAGAGCTTTATCAAG 360  
QY 424 CCAAGGCAAGAGTGAAGTTGGCACTGTGTATCACTTCACTTTTAAAGCTTTACAA 483  
DB 361 CCAAGGCAAGAGTGAAGTTGGCACTGTGTATCACTTCACTTTTAAAGCTTTTAAAG 420  
QY 484 ACCTCTTAAACATTTCTCATTAAGACCCCAATCAATTTGATCCAAAGTGTGTCA 543

DB 421 ACCTCTTAAACATTTCTCATTAAGACCCCAATCAATTTGATCCAAAGTGTGTCA 480  
QY 544 CAAGAAAGTGAATCTTGGAGTCAATTCGAAAATTGAGTATCTTCCCATCAATACT 603  
DB 481 CAAGAAAGTGAATCTTGGAGTCAATTCGAAAATTGAGTATCTTCCCATCAATACT 540  
QY 604 GGTGACTGTCTATTCAGAGTCAAGTGAATGACAGACATATTAATCAATTCAGGTT 663  
DB 541 GGTGACTGTCTATTCAGAGTCAAGTGAATGACAGACATATTAATCAATTCAGGTT 600  
QY 664 TCAGATATGTATTCAGAAAATTGAGTCACTTGGACAGACCAATTAATTTGTTATG 723  
DB 601 TCAGATATGTATTCAGAAAATTGAGTCACTTGGACAGACCAATTAATTTGTTATG 660  
QY 724 AATCTAAGCAATTTAAATGTACACATCAAGCAAGTATTAATTTGGAACCGAGTAA 783  
DB 661 AATCTAAGCAATTTAAATGTACACATCAAGCAAGTATTAATTTGGAACCGAGTAA 720  
QY 784 GGAGACGTAAAGCTTACATTTTACCTTATCTTTTGGGAAAGAAATTAATTA 843  
DB 721 GGAGACGTAAAGCTTACATTTTACCTTATCTTTTGGGAAAGAAATTAATTA 780  
QY 844 AAAACATTTAAGATTAATGATCTGCAAACTTCTTTTAAATGATGAAGATGAAAAT 903  
DB 781 AAAACATTTAAGATTAATGATCTGCAAACTTCTTTTAAATGATGAAGATGAAAAT 840  
QY 904 GTAATGATTTCTTCAAAATGAGACTTTTCAATACCTGATCTATCTTCCCTGACCACTA 963  
DB 841 GTAATGATTTCTTCAAAATGAGACTTTTCAATACCTGATCTATCTTCCCTGACCACTA 900  
QY 964 GAAATTTTAAACACAGTACAGATCACTTACAGTATTTTAAAGAAATGTAAGACTAAT 1023  
DB 901 GAAATTTTAAACACAGTACAGATCACTTACAGTATTTTAAAGAAATGTAAGACTAAT 960  
QY 1024 GTGTTCTTCAAGCAATGATTAATCACTTGAATTTTGAATTAATTAATTAATTAAT 1083  
DB 961 GTGTTCTTCAAGCAATGATTAATCACTTGAATTTTGAATTAATTAATTAATTAAT 1020  
QY 1084 CCATCTTCAACCTTCAACAGCCACTGTGAAAGTAACTGTGCTGATGGCAACCACTGACT 1143  
DB 1021 CCATCTTCAACCTTCAACAGCCACTGTGAAAGTAACTGTGCTGATGGCAACCACTGACT 1080  
QY 1144 CTGGAAGAAAGAAATTAATGTATGTATCACTAAGTACAGACAGAGAACTATTAAGAGTAC 1203  
DB 1081 CTGGAAGAAAGAAATTAATGTATGTATCACTAAGTACAGACAGAGAACTATTAAGAGTAC 1140  
QY 1204 TGAAGCGGATCTAAACAGTGAATTCAGAAATGGAAGCTGTTCAGAAAATTAATTAAT 1263  
DB 1141 TGAAGCGGATCTAAACAGTGAATTCAGAAATGGAAGCTGTTCAGAAAATTAATTAAT 1200  
QY 1264 GTCCCCCAAAGTGAACCTTTTAAATGAATTCATCCCAATCTGGAAGATTCAGAGAGCTA 1323  
DB 1201 GTCCCCCAAAGTGAACCTTTTAAATGAATTCATCCCAATCTGGAAGATTCAGAGAGCTA 1260  
QY 1324 CAGTTGAAGGCTTATTTCTTGTATGAATTAAGTGAAGTGAAGTCAAGTCTGTTTAAAG 1383  
DB 1261 CAGTTGAAGGCTTATTTCTTGTATGAATTAAGTGAAGTGAAGTCAAGTCTGTTTAAAG 1320  
QY 1384 TCTCTAAGTAAAGATCAATCACTAAACCAAGAGATGAAAATTAATTAAGGTTGAGTACG 1443  
DB 1321 TCTCTAAGTAAAGATCAATCACTAAACCAAGAGATGAAAATTAATTAAGGTTGAGTACG 1380  
QY 1444 CTTTGTAGTGTGTGTATGTGCAACAAAGATTAAGAGTGAAGTATTAATGTGTAAGT 1503  
DB 1381 CTTTGTAGTGTGTGTATGTGCAACAAAGATTAAGAGTGAAGTATTAATGTGTAAGT 1440  
QY 1504 TCCAGGGGACAGTGTGTGTGTGCAACAAAGATTAAGAGTGAAGTATTAATGTGTAAGT 1563  
DB 1441 TCCAGGGGACAGTGTGTGTGTGCAACAAAGATTAAGAGTGAAGTATTAATGTGTAAGT 1500  
QY 1564 GAAAATTTGTGACTCCAAAAGCTGTGTATTAATGTGTAAGTATTAATGTGTAAGTATGAGGAG 1623

Db 1501 GAAAAATCTTGGACTCCAAAAGCCTGTGTATTTGTATTTATTTAGATGATGGGAA 1560  
Qy 1524 ATTAATAGTATGTTCTTAAAAATTCCTGTTGAGCTTGTTTTAAATAAAGATAAGCTA 1683  
Db 1561 ATTATTAATGATGTTCTTAAAAATTCCTGTTGAGCTTGTTTTAAATAAAGATAAGCTA 1620  
Qy 1684 TATTGAGTAAAGTAAAGCTGAAACATCTGAGAAAGCTCTCTTTAGAGATCTGTGACA 1743  
Db 1621 TATTGAGTAAAGTAAAGCTGAAACATCTGAGAAAGCTCTCTTTAGAGATCTGTGACA 1680  
Qy 1744 CAGGCTGATCCGATAGTTGGATTTGTAGCTTTGACAAAAGTGGATCTGATGAATGCC 1803  
Db 1681 CAGGCTGATCCGATAGTTGGATTTGTAGCTTTGACAAAAGTGGATCTGATGAATGCC 1740  
Qy 1804 TCTATGATATTTACAAATGGAATGTGGTCCATGAGTTGAACTTTATTAACAGAGATAT 1863  
Db 1741 TCTATGATATTTACAAATGGAATGTGGTCCATGAGTTGAACTTTATTAACAGAGATAT 1800  
Qy 1864 TATTAGGCAATGTTCAATGATTTCTTTTGGAGTCTTTTCAGAAATGTGACCTGTGGATATG 1923  
Db 1801 TATTAGGCAATGTTCAATGATTTCTTTTGGAGTCTTTTCAGAAATGTGACCTGTGGATATG 1860  
Qy 1924 ACAGATGCAAACTTCAGAGAGATTTATTTGATGCTTTTATGCAATGCAAGATATGCT 1983  
Db 1861 ACAGATGCAAACTTCAGAGAGATTTATTTGATGCTTTTATGCAATGCAAGATATGCT 1920  
Qy 1984 GAGAGGTTTATGAGAAATGGAAGACATATTTGATATTTCACTTCTTTTGGGCT 2043  
Db 1921 GAGAGGTTTATGAGAAATGGAAGACATATTTGATATTTCACTTCTTTTGGGCT 1980  
Qy 2044 AGCAGTCCACATGTCGGAAGCATTTTTCAGAGACTTTGATTTGGCTAGACACCAACATG 2103  
Db 1981 AGCAGTCCACATGTCGGAAGCATTTTTCAGAGACTTTGATTTGGCTAGACACCAACATG 2040  
Qy 2104 GGTATACAGATTTTACCAAGATTTGAAAGTAACTGTAATCTGATTTCACTTCTTTGGGCT 2163  
Db 2041 GGTATACAGATTTTACCAAGATTTGAAAGTAACTGTAATCTGATTTCACTTCTTTGGGCT 2100  
Qy 2164 GCTACTGGTTTGTGATCTCTGAGGACTCTGGGCTTGGACTTAACAACTACTCCAGTGGAG 2223  
Db 2101 GCTACTGGTTTGTGATCTCTGAGGACTCTGGGCTTGGACTTAACAACTACTCCAGTGGAG 2160  
Qy 2224 CTCGAAGCTTTCACACATTTTTCATTTTGTGAATCTTCCCTACTCTGTATCAGAGCT 2283  
Db 2161 CTCGAAGCTTTCACACATTTTTCATTTTGTGAATCTTCCCTACTCTGTATCAGAGCT 2220  
Qy 2284 GAAAGATTTGCTTTGGAAATTAATCTATATTTCAATTTTGAAGAATGCCATGAGGTTAAG 2343  
Db 2221 GAAAGATTTGCTTTGGAAATTAATCTATATTTCAATTTTGAAGAATGCCATGAGGTTAAG 2280  
Qy 2344 GTATTCATTTGAGAAAGTGAACAAATTTGATTTCTAATGACTTCAAGTGAATTAATGCC 2403  
Db 2281 GTATTCATTTGAGAAAGTGAACAAATTTGATTTCTAATGACTTCAAGTGAATTAATGCC 2340  
Qy 2404 ACAGGCCACAGAGAGCCCTTCTGGTTCAGTGAAGATGGGCAACTGTTCTTTTCCC 2463  
Db 2341 ACAGGCCACAGAGAGCCCTTCTGGTTCAGTGAAGATGGGCAACTGTTCTTTTCCC 2400  
Qy 2464 ATCAGGCCAACATCTGGGAGAAATTCCTATCAAGTCAAGCTCTTTCAACCACTGCT 2523  
Db 2401 ATCAGGCCAACATCTGGGAGAAATTCCTATCAAGTCAAGCTCTTTCAACCACTGCT 2460  
Qy 2524 TCTGATGCTGCACCCAGATATTTTATTAAGGCTGAAGGAATGAAGAAATCAATTTCA 2583  
Db 2461 TCTGATGCTGCACCCAGATATTTTATTAAGGCTGAAGGAATGAAGAAATCAATTTCA 2520  
Qy 2584 CAATTCATCTTATTAAGCTTGAACAATAGGCTTACAGATACCTCTGAAAATCTTTGAGT 2643  
Db 2521 CAATTCATCTTATTAAGCTTGAACAATAGGCTTACAGATACCTCTGAAAATCTTTGAGT 2580  
Qy 2644 TTTCATATTTCTCTTAATACAGTGACTGGCAGTGAAGAGTTTCAGATCACTGCAATTTGA 2703  
Db 2581 TTTCATATTTCTCTTAATACAGTGACTGGCAGTGAAGAGTTTCAGATCACTGCAATTTGA 2640

Qy 2704 GATGTTCTTGCTGCTTCATCAATGAGCTTAGCCCTCATGATTTGAGATGCTTAGCTGT 2763  
Db 2641 GATGTTCTTGCTGCTTCATCAATGAGCTTAGCCCTCATGATTTGAGATGCTTAGCTGT 2700  
Qy 2764 GGTGAACAGAACATGATTAATTTTGTCTCCAAATATTTTCACTTTTGGATTTATGACTTAA 2823  
Db 2701 GGTGAACAGAACATGATTAATTTTGTCTCCAAATATTTTCACTTTTGGATTTATGACTTAA 2760  
Qy 2824 AAGAAACATGACAGATATTTGAAGAAAGAAAGCTCTTTCATTTATGAGGCAAGGTTAC 2883  
Db 2761 AAGAAACATGACAGATATTTGAAGAAAGAAAGCTCTTTCATTTATGAGGCAAGGTTAC 2820  
Qy 2884 CAGAGAACTTCTCTATCAGAGGAAAGATGCTCTTTCAGTGCCTTTTGGAAATTAAG 2943  
Db 2821 CAGAGAACTTCTCTATCAGAGGAAAGATGCTCTTTCAGTGCCTTTTGGAAATTAAG 2880  
Qy 2944 CTTCTGGAGACCTTGTTGTGAGCTTTTGTTTTAAATGATTTTCTTTGAAGCCGATCCT 3003  
Db 2881 CTTCTGGAGACCTTGTTGTGAGCTTTTGTTTTAAATGATTTTCTTTGAAGCCGATCCT 2940  
Qy 3004 TACATGATTTGATGAGAAATGTTTACACAGAACATACACTTGGCTTAAAGGACATCAG 3063  
Db 2941 TACATGATTTGATGAGAAATGTTTACACAGAACATACACTTGGCTTAAAGGACATCAG 3000  
Qy 3064 AAATCCAAAGGATTTTGGGATCCAGGAAGATGATTCATATGATGAGCTTCAAGGTGCG 3123  
Db 3001 AAATCCAAAGGATTTTGGGATCCAGGAAGATGATTCATATGATGAGCTTCAAGGTGCG 3060  
Qy 3124 AATTAAGTCCAGTAACACTTACAGCTTATTTGTAACTTCTCTCTGGGATTTAAGAA 3183  
Db 3061 AATTAAGTCCAGTAACACTTACAGCTTATTTGTAACTTCTCTCTGGGATTTAAGAA 3120  
Qy 3184 TATCAGGCTTACATGATGTCAGAGGTCATCTTTTGTGAGTGTGAATCAGTACA 3243  
Db 3121 TATCAGGCTTACATGATGTCAGAGGTCATCTTTTGTGAGTGTGAATCAGTACA 3180  
Qy 3244 GGAATTTACAGACATTTACTAGGCTTATTAATTAATGATGATGATGATGAGGAGT 3303  
Db 3181 GGAATTTACAGACATTTACTAGGCTTATTAATTAATGATGATGATGATGAGGAGT 3240  
Qy 3304 CTTAAAGCAGGAAGACTTGAATATGTCGACTTTGAGAGCAGAACAAAGGTGGCATG 3363  
Db 3241 CTTAAAGCAGGAAGACTTGAATATGTCGACTTTGAGAGCAGAACAAAGGTGGCATG 3300  
Qy 3364 CAATTCGGGTGCATCAGATGTCGAAACCTTTCGATCTCTGGAGCAGGCTCCCTGAT 3423  
Db 3301 CAATTCGGGTGCATCAGATGTCGAAACCTTTCGATCTCTGGAGCAGGCTCCCTGAT 3360  
Qy 3424 ATTGAATTCAGACCTATGACAGTCTCTCACACTTCTTACAAATTTGACATTTGAGGGA 3483  
Db 3361 ATTGAATTCAGACCTATGACAGTCTCTCACACTTCTTACAAATTTGACATTTGAGGGA 3420  
Qy 3484 ATCCCAATTAAGGTGGCTTAAGCAGGCAAGAAATGCTTGGGTGTTTGCATCTACT 3543  
Db 3421 ATCCCAATTAAGGTGGCTTAAGCAGGCAAGAAATGCTTGGGTGTTTGCATCTACT 3480  
Qy 3544 CAGGATTCACCTGGCTTTAAAGGCTCTGTCTGAATTTGACCCCTTAATGATACAGAA 3603  
Db 3481 CAGGATTCACCTGGCTTTAAAGGCTCTGTCTGAATTTGACCCCTTAATGATACAGAA 3540  
Qy 3604 AGGACAAATATCAAGTGAACGTCGAGGAGCTTACGCTCAACCAAGTCTCTTGGCTGTGTA 3663  
Db 3541 AGGACAAATATCAAGTGAACGTCGAGGAGCTTACGCTCAACCAAGTCTCTTGGCTGTGTA 3600  
Qy 3664 CAGCCAAATGCAAGTATTAATTTCCGAAATGTTTTGATTTGCTATTTGTACGCTCAAT 3723  
Db 3601 CAGCCAAATGCAAGTATTAATTTCCGAAATGTTTTGATTTGCTATTTGTACGCTCAAT 3660  
Qy 3724 GTTGTATTAATGTAAGGCTTCTGGGCTTCTTGAAGAAGAGATTTATCCAAATCA 3783  
Db 3661 GTTGTATTAATGTAAGGCTTCTGGGCTTCTTGAAGAAGAGATTTATCCAAATCA 3720

QY	3784	GAACCTTTGATTAGATGTGCTGTAAAGAAAATAAGATGATCTCAATCATGTGAT	3843
Db	3721	GAACCTTTGATTTAGATGTGCTGTAAAGAAAATAAGATGATCTCAATCATGTGAT	3780
QY	3844	TTGAATGTGTACAAAGCTTTTGGGGCCCGGGTAGAGATGSCATGCTCTTATGAAATT	3903
Db	3781	TTGAATGTGTACAAAGCTTTTGGGGCCCGGGTAGAGATGSCATGCTCTTATGAAATT	3840
QY	3904	AACTATTAATGCGCTTATATGTCCTTCAAGCAATTTCTCGAGGACACAGTAGAG	3963
Db	3841	AACTATTAATGCGCTTATATGTCCTTCAAGCAATTTCTCGAGGACACAGTAGAG	3900
QY	3964	AAAGTGAATATGATTCATGAGAAAACCTCAACTATTAGATTCGTATATGAACCAG	4023
Db	3901	AAAGTGAATATGATTCATGAGAAAACCTCAACTATTAGATTCGTATATGAACCAG	3960
QY	4024	TTTTTGTTAAATATCTCGCTGTGAGAACTTTAAAGTTCAATAATCCCAAGATGCTTCA	4083
Db	3961	TTTTTGTTAAATATCTCGCTGTGAGAACTTTAAAGTTCAATAATCCCAAGATGCTTCA	4020
QY	4084	GTGTCCATAGTGAATTACTATAGGACAAAGAGACAGGCGGTGAGAAAGTTACAATCTGAA	4143
Db	4021	GTGTCCATAGTGAATTACTATAGGACAAAGAGACAGGCGGTGAGAAAGTTACAATCTGAA	4080
QY	4144	GTGAAGCTGTCTCTGTGACCTTTGCAGTAGATGTCCAGGGCTGCCGCTCTTGTGAGAT	4203
Db	4081	GTGAAGCTGTCTCTGTGACCTTTGCAGTAGATGTCCAGGGCTGCCGCTCTTGTGAGAT	4140
QY	4204	GGAGCT 4209	
Db	4141	GGAGCT 4146	

ID	AAID49437	standard; DNA; 4197 BP.
AC	AAID49437;	
DT	24-MAR-2003	(first entry)
DE	Human r150 DNA #2.	
KM	Human; GPI-anchored TGF-beta1 binding protein; r150 protein; cancer	
KM	glycosylphosphatidyl inositol; transforming growth factor-beta1;	
KM	therapy; gene; ds.	
OS	Homo sapiens.	
FT	Key	Location/Qualifiers
FT	CDS	1..4197
FT	/*tag= a	/product= "Human r1520 protein #3"
FT	/transl_except= (pos:2044..2046, aa:Xaa)	
FT	/note= "Xaa corresponds to Ser, Tyr; No start	
FT	and stop codon"	
FT	/partial	
PN	WO200285942-A2.	
PD	31-OCT-2002.	
PF	24-APR-2002; 2002WO-CA00560.	
PR	24-APR-2001; 2001US-285713P.	
PR	14-FEB-2002; 2002US-356163P.	
PA	(UYMC-) UNIV MCGILL.	
PI	Philip A, Tam B;	
WP	WPI; 2003-093100/08.	
DR	P-PSDB; AAB32015.	

XX Novel transforming growth factor (TGF)-beta 1 binding reagent which  
PT comprises r150 protein which acts as accessory receptor of TGF-beta,  
PT useful for negatively modulating TGF-beta activity, and thus for  
PT treating cancer -  
XX  
PS  
PS Claim 7; Page 118-120; 127bp; English.  
XX  
XX The invention relates to novel transforming growth factor (TGF)-beta1  
CC binding reagent which comprises a glycosylphosphatidylinositol (GPI)-  
CC anchored TGF-beta1 binding protein referred to as r150 which acts as  
CC an accessory receptor of TGF-beta. The invention is used for negatively  
CC modulating TGF-beta activity, and thus for treating conditions  
CC characterised by overproduction of TGF-beta, such as cancer. Antisense  
CC molecules of the invention are used for increasing TGF-beta availability  
CC and increase graft success. The present sequence is human r150 protein  
CC encoding DNA.  
XX  
XX Sequence 4197 BP; 1276 A; 788 C; 894 G; 1238 T; 1 other;  
SQ

Query Match	95.2%	Score 4082.4	DB 25	Length 4197
Best Local Similarity	98.7%	Pred. No. 0		
Matches 144	Conservative	0	Mismatch	2
			Indels	51
			Gaps	1
Qy	64	GTGGCTCCCGGGCTCGGGTTCTGGTGAACAGCCCCAGGSAATCATCAGGCCCCGAGGAAT	123	
Db	1	GTGGCTCCCGGGCTCGGGTTCTGGTGAACAGCCCCAGGSAATCATCAGGCCCCGAGGAAT	60	
Qy	124	GTGACTATTGGGGGTGAGAGCTTCTGGAAACATGCCCCCTTACAGGTGACTGTGAAGGCGAG	183	
Db	61	GTGACTATTGGGGGTGAGAGCTTCTGGAAACATGCCCCCTTACAGGTGACTGTGAAGGCGAG	120	
Qy	184	CTGCTCAAGACAGCATCAAACTCAGTCTCTGTGCTCGAAGCAGAAGAGTCTTTGAA	243	
Db	121	CTGCTCAAGACAGCATCAAACTCAGTCTCTGTGCTCGAAGCAGAAGAGTCTTTGAA	180	
Qy	244	AAAGGCTCTTTAAGACATTAATCTTTCATCACTACCTCTGAACAGTGCAATGAT	303	
Db	181	AAAGGCTCTTTAAGACATTAATCTTTCATCACTACCTCTGAACAGTGCAATGAT	240	
Qy	304	TATGAGCTAGCTGTAAACCGGACGTACCCAGATGAGATTTATTTCTTAATAGTACCGC	363	
Db	241	TATGAGCTAGCTGTAAACCGGACGTACCCAGATGAGATTTATTTCTTAATAGTACCGC	300	
Qy	364	TTATCATTTTGGAGCCAGAGAAATCTGCTCTCATTTCAACAGACAGGCGCTTATACAG	423	
Db	301	TTATCATTTTGGAGCCAGAGAAATCTGCTCTCATTTCAACAGACAGGCGCTTATACAG	360	
Qy	424	CCAAAGCAGAAGGAGAGTTTTCGATTTGTTAACTCTTTCAGATTTTAAAGCTTACAA	483	
Db	361	CCAAAGCAGAAGGAGAGTTTTCGATTTGTTAACTCTTTCAGATTTTAAAGCTTACAA	420	
Qy	484	ACCTCTTTAAACATCTTCATTAAAGACCCCAATCAAAATTTGATCAACAGTGGTTGCA	543	
Db	421	ACCTCTTTAAACATCTTCATTAAAGACCCCAATCAAAATTTGATCAACAGTGGTTGCA	480	
Qy	544	CAACAAAGTATCTTGGAGTCAATTCGAAACCTTTCAAGCTATCTTCCCATCAATACTT	603	
Db	481	CAACAAAGTATCTTGGAGTCAATTCGAAACCTTTCAAGCTATCTTCCCATCAATACTT	540	
Qy	604	GGTGAATGTCATATTCAGTTCAAGTTCAAGTGAACGACAGCATTTATTCATCATCTTCAAGTT	663	
Db	541	GGTGAATGTCATATTCAGTTCAAGTTCAAGTGAACGACAGCATTTATTCATCATCTTCAAGTT	600	
Qy	664	TCAGAAATATGTAATTCACAAATTTGAAGTGACTTTGGAGACACCATTAATATTTGTTCTATG	723	
Db	601	TCAGAAATATGTAATTCACAAATTTGAAGTGACTTTGGAGACACCATTAATATTTGTTCTATG	660	
Qy	724	AATTTCAAGCATTTTAAATGTAATGTAACCATCAGGCAAAAGTATCATATATGGAAAGCCAGTGA	783	
Db	661	AATTTCAAGCATTTTAAATGTAATGTAACCATCAGGCAAAAGTATCATATATGGAAAGCCAGTGA	720	
Qy	784	GGAGACGTAAAGCTTACATTTTAACTTTTATCTTTTGGGAAAGAAAGAAATATATTACA	843	

```
Db 721 GGAGACGTACGCTTACATTTTACCTTTATCTTTGGGAGAAAGAAAATATTACA 780
Qy 844 AAAACATTAAAGTAATGAGATCGCAAACTTCCTTTTAAATGATGAGATGAAAAAT 903
Db 781 AAAACATTAAAGTAATGAGATCGCAAACTTCCTTTTAAATGATGAGATGAAAAAT 840
Qy 904 GTAATGATCTTCAATGAGATCTTCTGATATCTGATCTATCTTCCCTGAGCACTA 963
Db 841 GTATGATCTTCAATGAGATCTTCTGATATCTGATCTATCTTCCCTGAGCACTA 900
Qy 964 GAAATTTTAAACAGTACAGATCAAGTATCAAGTATTTCAAGAAATGTAAGACTAT 1023
Db 901 GAAATTTTAAACAGTACAGATCAAGTATCAAGTATTTCAAGAAATGTAAGACTAT 960
Qy 1024 GTGTTCTTCAAGCAACATGATTTACATCATGATTTTGTATATATCTGATGAG 1083
Db 961 GTGTTCTTCAAGCAACATGATTTACATCATGATTTTGTATATATCTGATGAG 1020
Qy 1084 CCATCTCTCAACTTCAACAGCACTGTGAGGTAATCTGCTGATGAGCAACCACTGACT 1143
Db 1021 CCATCTCTCAACTTCAACAGCACTGTGAGGTAATCTGCTGATGAGCAACCACTGACT 1080
Qy 1144 CTGAGAGAAAGAAATATATGATCTATTAACAGTACACAGAGAACTATATGAGTAC 1203
Db 1081 CTGAGAGAAAGAAATATATGATCTATTAACAGTACACAGAGAACTATATGAGTAC 1140
Qy 1204 TGGAGCGGATCTTAACAGTGAATATCAAGAAATGAGAGTCTTCAAGAAATTAATAT 1263
Db 1141 TGGAGCGGATCTTAACAGTGAATATCAAGAAATGAGAGTCTTCAAGAAATTAATAT 1200
Qy 1264 GTCCCCCAAGTGAACCTTTTAAAGTGAATTTCCCAATCTCGAGGATTCAGTGAAGTA 1323
Db 1201 GTCCCCCAAGTGAACCTTTTAAAGTGAATTTCCCAATCTCGAGGATTCAGTGAAGTA 1260
Qy 1324 CAGTTGAAGGCTTATTTCTTGTATGAATAAGTACATGAGTCTGTTTAAAG 1383
Db 1261 CAGTTGAAGGCTTATTTCTTGTATGAATAAGTACATGAGTCTGTTTAAAG 1320
Qy 1384 TCTCTCTAGTACATCATCACTAATAAACAAGATGAGAAATTTAAAGTGGATCG 1443
Db 1321 TCTCTCTAGTACATCATCACTAATAAACAAGATGAGAAATTTAAAGTGGATCG 1380
Qy 1444 CCTTTGAGTGGTGTAGTGGCAACAAGATGAGAGATTAAGCTATATGATGATG 1503
Db 1381 CCTTTGAGTGGTGTAGTGGCAACAAGATGAGAGATTAAGCTATATGATGATG 1440
Qy 1504 TCCAGGGGACAGTTGGTGGCTGTAGAGAAACAAAATTTCAACATGTTCTTTTAA 1563
Db 1441 TCCAGGGGACAGTTGGTGGCTGTAGAGAAACAAAATTTCAACATGTTCTTTTAA 1500
Qy 1564 GAAATTTTGGAGCTCCAAAAGCTGTATTTGTATTTATTTGAAGATGAGGGAA 1623
Db 1501 GAAATTTTGGAGCTCCAAAAGCTGTATTTGTATTTATTTGAAGATGAGGGAA 1560
Qy 1624 ATTAATGATGATTTCTTAAATTTCTGTTGAGCTGTTTAAATTAATTAAGTAA 1683
Db 1561 ATTAATGATGATTTCTTAAATTTCTGTTGAGCTGTTTAAATTAATTAAGTAA 1620
Qy 1684 TATTGAGTAAAGTAAAGTGAACCATCTGAGAAAGTCTCTTGAAGATCTGTGACA 1743
Db 1621 TATTGAGTAAAGTAAAGTGAACCATCTGAGAAAGTCTCTTGAAGATCTGTGACA 1680
Qy 1744 CAGCTGACTCCATATGTTGGAGTTGAGCTTTGACAAAGTGTGATCTGATGATGCC 1803
Db 1681 CAGCTGACTCCATATGTTGGAGTTGAGCTTTGACAAAGTGTGATCTGATGATGCC 1740
Qy 1804 TCTATATGATTTTAAACAATGGAATGTGCTCATGATGTTGAACCTTTTAAACAAGAT 1863
Db 1741 TCTATATGATTTTAAACAATGGAATGTGCTCATGATGTTGAACCTTTTAAACAAGAT 1800
Qy 1864 TATTAGGATGTTCAATGATTTCTTTGAGTCTTTCAAGAAATGAGACTCTGGGATTTG 1923
Db 1801 TATTAGGATGTTCAATGATTTCTTTGAGTCTTTCAAGAAATGAGACTCTGGGATTTG 1860
Qy 1924 ACAGATGCAAAACCTCAGAAAGATTTATGATGATGATGATGATGATGATGATGATG 1983
Db 1861 ACAGATGCAAAACCTCAGAAAGATTTATGATGATGATGATGATGATGATGATGATG 1920
Qy 1984 GAGAGGTTTATGAGAGAAATGAGAGCATATTTGATATATTTGATGATGATGATGATG 2043
Db 1921 GAGAGGTTTATGAGAGAAATGAGAGCATATTTGATATATTTGATGATGATGATGATG 1980
Qy 2044 AGCAGTCCATGATCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAATG 2103
Db 1981 AGCAGTCCATGATCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAATG 2040
Qy 2104 GGTTCAGATTTTACCAAGATTTGAGATGATGATGATGATGATGATGATGATGATG 2163
Db 2041 GGTTCAGATTTTACCAAGATTTGAGATGATGATGATGATGATGATGATGATGATG 2100
Qy 2164 GCTACTGTTTGTGATCTCTGAGAGACTGGGCTTTGAGACTTAAACAACCTGAGTGA 2223
Db 2101 GCTACTGTTTGTGATCTCTGAGAGACTGGGCTTTGAGACTTAAACAACCTGAGTGA 2160
Qy 2224 CTGAGAGCTTCCAGACATTTTCAATTTTGTGATTTTGTGATTTTGTGATTTTGTG 2283
Db 2161 CTGAGAGCTTCCAGACATTTTCAATTTTGTGATTTTGTGATTTTGTGATTTTGTG 2220
Qy 2284 GAGAAATTTGCTTGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2343
Db 2221 GAGAAATTTGCTTGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2280
Qy 2344 GTATCATGAGAGAAAGTGAACAATTTGATATTTGATGATTTCAAGTGAATTAATGCC 2403
Db 2281 GTATCATGAGAGAAAGTGAACAATTTGATATTTGATGATTTCAATGATTAATTAATGCC 2340
Qy 2404 ACAGGCCACAGAGACCTTTCTGTTCCAGTGAAGATGAGGGCAACTGTTCTTTTCCC 2463
Db 2341 ACAGGCCACAGAGACCTTTCTGTTCCAGTGAAGATGAGGGCAACTGTTCTTTTCCC 2400
Qy 2464 ATCAGGCCAACATCTGGGAGAAATTTCCATGACAGTCAACAGCTCTTCAACCACTGCT 2523
Db 2401 ATCAGGCCAACATCTGGGAGAAATTTCCATGACAGTCAACAGCTCTTCAACCACTGCT 2460
Qy 2524 TCTGATGCTGTCAACCAAGATTTTGAATAAGGCTGAAGAAATGAGAAATCATATTTCA 2583
Db 2461 TCTGATGCTGTCAACCAAGATTTTGAATAAGGCTGAAGAAATGAGAAATCATATTTCA 2520
Qy 2584 CAATCATCTTATTAAGCTTGAAGTCAATGAGCTCAAGATGACCTGAAACCTTTGAGT 2643
Db 2521 CAATCATCTTATTAAGCTTGAAGTCAATGAGCTCAAGATGACCTGAAACCTTTGAGT 2580
Qy 2644 TTCTCATTTCTCTTAATTAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 2703
Db 2581 TTCTCATTTCTCTTAATTAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 2640
Qy 2704 GATGTTCTTGGTCTTCAATCAATGAGCTTGAAGCTCAATGATGATGATGATGATGATG 2763
Db 2641 GATGTTCTTGGTCTTCAATCAATGAGCTTGAAGCTCAATGATGATGATGATGATGATG 2700
Qy 2764 GGTGAACAGAACATGATTAATTTTGTCCAAATATTTTAAATTTTGAATTAATGATG 2823
Db 2701 GGTGAACAGAACATGATTAATTTTGTCCAAATATTTTAAATTTTGAATTAATGATG 2760
Qy 2824 AAGAAACAACGACATATTAATTTGAAGAAAGCTCTTCAATTTTGAAGGCAAGTTAC 2883
Db 2761 AAGAAACAACGACATATTAATTTGAAGAAAGCTCTTCAATTTTGAAGGCAAGTTAC 2820
Qy 2884 CAGAGAACTTCTCATGAGAGAGATGAGCTCTTCAAGTCTTGGGAAATTTGATGAC 2943
Db 2821 CAGAGAACTTCTCATGAGAGAGATGAGCTCTTCAAGTCTTGGGAAATTTGATGAC 2880
Qy 2944 CTTTCTGGAGACCTTGGTTGAGCTTTTGAATGATTTTCTTGAAGCTGATCTT 3003
Db 2881 CTTTCTGGAGACCTTGGTTGAGCTTTTGAATGATTTTCTTGAAGCTGATCTT 2940
```

QY 3004 TACATGATATTTGATCAGATGTGTTACACAGAACATACCTTGCTTAAGACATCAG 3063  
 DB 2941 TACATGATATTTGATCAGATGTGTTACACAGAACATACCTTGCTTAAGACATCAG 3000  
 QY 3064 AATATCCAGGGTGAATTTGGATCCAGAGAGATGATTCATATGTAGCTTCAAGGTGC 3123  
 DB 3001 AATATCCAGGGTGAATTTGGATCCAGAGAGATGATTCATATGTAGCTTCAAGGTGC 3060  
 QY 3124 AATTAAGTCCAGTACCTTACAGCTATATTTGATCTCTCTCTGGATATTAAG 3183  
 DB 3061 AATTAAGTCCAGTACCTTACAGCTATATTTGATCTCTCTCTGGATATTAAG 3120  
 QY 3184 TATCAGCCTTAACATTTGATGAGAGAGATTCATCTATTTTGGAGCTGAATTCAGTAG 3243  
 DB 3121 TATCAGCCTTAACATTTGATGAGAGAGATTCATCTATTTTGGAGCTGAATTCAGTAG 3180  
 QY 3244 GGAATTTCAAGCAATTAATCTTACGCTTATTAATTAATGATGATGATGAGGAGT 3303  
 DB 3181 GGAATTTCAAGCAATTAATCTTACGCTTATTAATTAATGATGATGATGAGGAGT 3240  
 QY 3304 CCTAAGGAGAGAGCTTTGAATATGCTGATGAGAGAGCAACAAAGGTGAGT 3363  
 DB 3241 CCTAAGGAGAGAGCTTTGAATATGCTGATGAGAGAGCAACAAAGGTGAGT 3300  
 QY 3364 CAATTCGGGTGTCATCAGATGCCAACTTTGATCTCTGGAGAGCAGCTCCCTGAT 3423  
 DB 3301 CAATTCGGGTGTCATCAGATGCCAACTTTGATCTCTGGAGAGCAGCTCCCTGAT 3360  
 QY 3424 ATTTGAAGTTCAGCTTACCTGCTCTGACACTTCTTCAATTTGACACTTCTGAGGGA 3483  
 DB 3361 ATTTGAAGTTCAGCTTACCTGCTCTGACACTTCTTCAATTTGACACTTCTGAGGGA 3420  
 QY 3484 ATTTGAAGTTCAGCTTACCTGCTCTGACACTTCTTCAATTTGACACTTCTGAGGGA 3543  
 DB 3421 ATTTGAAGTTCAGCTTACCTGCTCTGACACTTCTTCAATTTGACACTTCTGAGGGA 3480  
 QY 3544 CAGATATACCATGCTGCTTTAAAGGCTCTGATTTGACAGCTTATGATGATGATGAT 3603  
 DB 3481 CAGATATACCATGCTGCTTTAAAGGCTCTGATTTGACAGCTTATGATGATGATGAT 3540  
 QY 3604 AGGACAAATATTCAGAGTGAACGCTGACGGGCTTACCTCAAGTCTCTCTCTCTCT 3651  
 DB 3541 AGGACAAATATTCAGAGTGAACGCTGACGGGCTTACCTCAAGTCTCTCTCTCTCT 3600  
 QY 3652 -----CTTCT 3672  
 DB 3601 ATTTGACACACACACGCTTACTCTTCAAGACAGACAGACCTTCTCTCTCTCTCTCT 3660  
 QY 3673 GCAGTTAATATTTCCGCAATGATTTGATTTGCTATTTGCTATTTGCTATTTGCTATTT 3732  
 DB 3661 GCAGTTAATATTTCCGCAATGATTTGATTTGCTATTTGCTATTTGCTATTTGCTATTT 3720  
 QY 3733 AATGTGAAGGCTTCTGAGTCTCTGAGAGACAGAGATCTTATCCAAATCAAGAGCTTT 3792  
 DB 3721 AATGTGAAGGCTTCTGAGTCTCTGAGAGACAGAGATCTTATCCAAATCAAGAGCTTT 3780  
 QY 3793 GATTTAGATGTTGCTGTTAAAGAAATTAAGATGATCTCAATCATGATGATTTGATG 3852  
 DB 3781 GATTTAGATGTTGCTGTTAAAGAAATTAAGATGATCTCAATCATGATGATTTGATG 3840  
 QY 3853 TGTACAGCTTTCCGGGCGCGGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3912  
 DB 3841 TGTACAGCTTTCCGGGCGCGGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3900  
 QY 3913 AGTGGCTTATGATGCTTCAAGAGCAATTTCTGAGAGAGAGAGAGAGAGAGAGAGAG 3972  
 DB 3901 AGTGGCTTATGATGCTTCAAGAGCAATTTCTGAGAGAGAGAGAGAGAGAGAGAGAG 3960  
 QY 3973 TATGATCATGAGAAACTCAACTCTATTAGATTTCTGTAATGAAACCAAGTTTGTGTT 4032  
 DB 3961 TATGATCATGAGAAACTCAACTCTATTAGATTTCTGTAATGAAACCAAGTTTGTGTT 4020

QY 4033 AATATTCCTGCTGTGAGAACTTTAAAGTTTCAATATCCAAATGCTTCAATGATCATA 4092  
 DB 4021 AATATTCCTGCTGTGAGAACTTTAAAGTTTCAATATCCAAATGCTTCAATGATCATA 4080  
 QY 4093 GTGATTTCTATGAGCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4152  
 DB 4081 GTGATTTCTATGAGCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4140  
 QY 4153 TCTCTCTGAGCTTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 4209  
 DB 4141 TCTCTCTGAGCTTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 4197  
 RESULT 14  
 AB079968  
 ID AB079968 standard; cDNA, 3535 BP.  
 XX  
 AC AB079968;  
 XX  
 DT 23-DEC-2002 (first entry)  
 XX  
 DE Human CD109 K15 protein encoding cDNA.  
 XX  
 KW CD109; alpha2 macroglobulin; thioester; cerebroprotective; cardiant;  
 KW immunosuppressive; haemostatic; anticoagulant; thrombolytic; human;  
 KW cardiovascular; vasotropic; gene therapy; CD109 K15; gene; ss.  
 OS Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FH 113..4450  
 FT /tag= a  
 FT /product= "CD109 K15"  
 XX  
 PN WO200270696-A2.  
 XX  
 PD 12-SEP-2002.  
 XX  
 PR 07-MAR-2002; 2002MO-CA00292.  
 XX  
 PR 07-MAR-2001; 2001US-273814P.  
 XX  
 PA (SCHU/) SCHUH A.  
 PA (SUTH/) SUTHERLAND R D.  
 XX  
 PI Schuh A, Sutherland RD;  
 XX  
 DR WPI, 2002-713450/77.  
 DR P-PSDB; ABB82169.  
 XX  
 PT New CD109 nucleic acids and polypeptides, useful in gene therapy,  
 PT particularly for treating strokes, myocardial infarctions, thrombosis,  
 PT thrombocytopenia, autoimmune diseases, or organ or bone marrow  
 PT transplantation -  
 XX  
 PS Claim 1; Fig 4a; 156bp; English.  
 XX  
 CC The invention relates to isolated nucleic acid molecules encoding CD109  
 CC polypeptides. These nucleic acid molecules include the human cDNA  
 CC sequences comprising CD109 K1, CD109 K1-H7, CD109 K15 or their variants.  
 CC CD109 is a novel member of the alpha2 macroglobulin (alpha2M)/C3, C4, C5  
 CC family of thioester-containing proteins. The CD109 polypeptides can be  
 CC expressed by standard recombinant methodology. The CD109 nucleic acid,  
 CC CD109 polypeptide or its mimetic, or the anti-CD109 antibody is useful  
 CC for treating or detecting a disease or disorder, e.g. conditions  
 CC associated with endothelial activation, platelet activation, activation  
 CC of the coagulation or fibrinolytic systems, activation of T lymphocytes  
 CC and of the complement system, quantitative or qualitative abnormalities  
 CC of platelet function, increased or impaired platelet aggregation and  
 CC activation, increased or impaired activation of the coagulation and/or  
 CC fibrinolytic systems, or impaired or increased immune activation. These  
 CC are also useful for treating cardiovascular disorders, stroke, myocardial  
 CC infarction, thrombosis, embolism, peripheral vascular disease,

CC thrombocytopenia, thrombocythemia, autoimmune diseases, organ  
CC transplantation, or bone marrow transplantation. The CD109 nucleic acids  
CC or their antisense nucleotide sequence are useful in gene therapy for  
CC treating these conditions. The present sequence represents the human  
CC CD109 K15 cDNA sequence.

XX Sequence 3535 BP; 1108 A; 675 C; 731 G; 1021 T; 0 other;

Query Match 74.4%; Score 3187.4; DB 24; Length 3535;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3188; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAGAGCCCAACCGCTCTGACCGCGCCCACTCTCTGCTGCTGACCGCGCGCTG 60  
DB 113 ATGAGAGGCCCAACCGCTCTGACCGCGCCCACTCTCTGCTGCTGACCGCGCGCTG 172  
QY 61 GCCCGGCTCCCGGGCCCTGGGTTTCTGGTGAAGCCCGAGGATCATAGCCCGGAGGA 120  
DB 173 GCCCGGCTCCCGGGCCCTGGGTTTCTGGTGAAGCCCGAGGATCATAGCCCGGAGGA 232  
QY 121 AATGTGACTATTTGGGGTGGAGCTTCTGGAGACCTGCCCTTCAAGTGAAGGAGCG 180  
DB 233 AATGTGACTATTTGGGGTGGAGCTTCTGGAGACCTGCCCTTCAAGTGAAGGAGCG 292  
QY 181 GAGCTGCTCAAGACATCAACCTCACTGTCTCTGTCTGAGACGAGAGGCTCTT 240  
DB 293 GAGCTGCTCAAGACATCAACCTCACTGTCTCTGTCTGAGAGGAGGAGCTCTT 352  
QY 241 GAAAAAGGCTTTTAAAGACCTTACTCTTCATCACTCACTCGTGAAGAGGAGTGA 300  
DB 353 GAAAAAGGCTTTTAAAGACCTTACTCTTCATCACTCACTCGTGAAGAGGAGTGA 412  
QY 301 ATTATGAGTACGCTGTAAACCGAGCTACCGAGATGAGATTTTATCTTAATAGTACC 360  
DB 413 ATTATGAGTACGCTGTAAACCGAGCTACCGAGATGAGATTTTATCTTAATAGTACC 472  
QY 361 CGCTTATCATTTGAGACCAAGAGAAATATCTGTCTTCAATCAACAGACAGGCTTATAC 420  
DB 473 CGCTTATCATTTGAGACCAAGAGAAATATCTGTCTTCAATCAACAGACAGGCTTATAC 532  
QY 421 AAGCCAAAGCAAGAGAAATTTGCGATTTGTTCACCTCTCTGAGATTTTAAACCTTAC 480  
DB 533 AAGCCAAAGCAAGAGAAATTTGCGATTTGTTCACCTCTCTGAGATTTTAAACCTTAC 592  
QY 481 AAAACCTCTTAAACATCTCATTTAAGGACCCCAATCAATTTGATCCACAGTGTG 540  
DB 593 AAAACCTCTTAAACATCTCATTTAAGGACCCCAATCAATTTGATCCACAGTGTG 652  
QY 541 TCACAAACAAGTGAATCTTGAAGTCAATTCAAAACCTTTCAGCTATCTTCCATCAATA 600  
DB 653 TCACAAACAAGTGAATCTTGAAGTCAATTCAAAACCTTTCAGCTATCTTCCATCAATA 712  
QY 601 CTGTGTGACTGTGCTATTTCAAGTTCAGTGAATGACGACATTTATCATCTTTTAC 660  
DB 713 CTGTGTGACTGTGCTATTTCAAGTTCAGTGAATGACGACATTTATCATCTTTTAC 772  
QY 661 GTTTCAGAAATATGATTAACAAAATTTGAAGTGAATTTGACAGACCACTTATTTGTTCT 720  
DB 773 GTTTCAGAAATATGATTAACAAAATTTGAAGTGAATTTGACAGACCACTTATTTGTTCT 832  
QY 721 ATGAATTTAAGCATTTAATATGATGATCAACGAGCAAGATATATATGAGGAGCCAGTG 780  
DB 833 ATGAATTTAAGCATTTAATATGATGATCAACGAGCAAGATATATATGAGGAGCCAGTG 892  
QY 781 AAAGAGAGCTTAAGGCTTACATTTTACCTTTATCTTTTGGGAAAGAGAAAAATATT 840  
DB 893 AAAGAGAGCTTAAGGCTTACATTTTACCTTTATCTTTTGGGAAAGAGAAAAATATT 952  
QY 841 ACAAAAACATTATAGATTAATAGATTCGCAAACTTCTCTTTAATGAGAGAGATGAA 900  
DB 953 ACAAAAACATTATAGATTAATAGATTCGCAAACTTCTCTTTAATGAGAGAGATGAA 1012  
QY 901 AATGTAATGATCTTCAATATGAGATTTCTGAAATACCTGATCTATCTTCCCTGAGCA 960

DB 1013 AATGTAATGATCTTCAATATGAGATTTCTGAAATACCTGATATATCTTCCCTGAGCA 1072  
QY 961 GTAGAAATTTTAAACACAGTGAAGATCAGTATACAGTATTTCAAGAAATGATAGACT 1020  
DB 1073 GTAGAAATTTTAAACACAGTGAAGATCAGTATACAGTATTTCAAGAAATGATAGACT 1132  
QY 1021 AATGTGCTCTTCAAGCAACATGATATCATCATTTAGTTTGTATTAATTAATTAATTAAT 1080  
DB 1133 AATGTGCTCTTCAAGCAACATGATATCATCATTTAGTTTGTATTAATTAATTAATTAAT 1192  
QY 1081 AAGCATCTCTCAACTTCAAGCACTGTAAGTAACTCTGCTGATGAGCAACCACTG 1140  
DB 1193 AAGCATCTCTCAACTTCAAGCACTGTAAGTAACTCTGCTGATGAGCAACCACTG 1252  
QY 1141 ACTCTGAAAGAAAGAAATTAATGATATGATATCAATACAGTGAACAGAGAAATTAATCTAG 1200  
DB 1253 ACTCTGAAAGAAAGAAATTAATGATATGATATCAATACAGTGAACAGAGAAATTAATCTAG 1312  
QY 1201 TACTGAGCGGATCTAAACAGTGAATAATCAAGAAATGAGAGCTGTTCAGAAATTAATTAAT 1260  
DB 1313 TACTGAGCGGATCTAAACAGTGAATAATCAAGAAATGAGAGCTGTTCAGAAATTAATTAAT 1372  
QY 1261 ACTGTCCCCAAGTGAATTTTAAATGATTAATCCCAATCTCTGAGAGATTCAGAGTGA 1320  
DB 1373 ACTGTCCCCAAGTGAATTTTAAATGATTAATCCCAATCTCTGAGAGATTCAGAGTGA 1432  
QY 1321 CTACAGTTGAAGGCTATTTCTTGGTATGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1380  
DB 1433 CTACAGTTGAAGGCTATTTCTTGGTATGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1492  
QY 1381 AAGTCTCTTATGTAAGATATCATCATCTAAATCAAGAAATGAGATGAAATTAATTAATTAAT 1440  
DB 1493 AAGTCTCTTATGTAAGATATCATCATCTAAATCAAGAAATGAGATGAAATTAATTAATTAAT 1552  
QY 1441 TCGCTTTTGAATGTTGGTGTAGTGAAGCAAAAGATGAAGAGTTAAGCTATATGTA 1500  
DB 1553 TCGCTTTTGAATGTTGGTGTAGTGAAGCAAAAGATGAAGAGTTAAGCTATATGTA 1612  
QY 1501 GTATCCAGGGGACAGTGTGGTGTAGGAAACAAATTCACAAATGTTCTCTTTTACA 1560  
DB 1613 GTATCCAGGGGACAGTGTGGTGTAGGAAACAAATTCACAAATGTTCTCTTTTACA 1672  
QY 1561 CCAGAAATTTCTTGAATCCAAAGGCTGTGATTTGTATTAATTAAGATGATGG 1620  
DB 1673 CCAGAAATTTCTTGAATCCAAAGGCTGTGATTTGTATTAATTAAGATGATGG 1732  
QY 1621 GAAATTAAGTATGTTCTTAAATTTCTGTTCAGCTTGTTTTAAATTAAGATTAAG 1680  
DB 1733 GAAATTAAGTATGTTCTTAAATTTCTGTTCAGCTTGTTTTAAATTAAGATTAAG 1792  
QY 1681 CTATATTTGAGTAAAGTGAAGGCTGAACATCTGAGAAAGCTCTCTAGATCTCTGTG 1740  
DB 1793 CTATATTTGAGTAAAGTGAAGGCTGAACATCTGAGAAAGCTCTCTAGATCTCTGTG 1852  
QY 1741 ACAAGGCTGATCTCAATAGTTGGATTTGATCTGTGACAAAGGATGATGATGAAT 1800  
DB 1853 ACAAGGCTGATCTCAATAGTTGGATTTGATCTGTGACAAAGGATGATGATGAAT 1912  
QY 1801 GCTCTTAATGATATTAATGATGAAATGTGTCAATGATGATGATGATGATGATGATGATGAT 1860  
DB 1913 GCTCTTAATGATATTAATGATGAAATGTGTCAATGATGATGATGATGATGATGATGATGAT 1972  
QY 1861 TATTTATTTAGGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920  
DB 1973 TATTTATTTAGGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2032  
QY 1921 TTGACAGATCAACCTCAAGAGATTAATGATGATGATGATGATGATGATGATGATGATGAT 1980  
DB 2033 TTGACAGATCAACCTCAAGAGATTAATGATGATGATGATGATGATGATGATGATGATGAT 2092  
QY 1981 GCTGAGAGTTTATGAGAGAAATGAGACATATTTGTATGATATTCATGACTTTCTTTG 2040



Db 2093 GCTGAGAGGTTTATGAGAGAAAATGAGACATATTGTAGATATTCATGACTTTTCTTTG 2152  
 Qy 2041 GGTAGACGTCACATGTCCGAAAGCATTTTCCAGAGACTTGATTTGGCTAGACACCAAC 2100  
 Db 2153 GGTAGACGTCACATGTCCGAAAGCATTTTCCAGAGACTTGATTTGGCTAGACACCAAC 2212  
 Qy 2101 ATGGGTACAGAGATTTACCAAGAAATTTGAAGTAACTGTAAGTATTTGATTCATCTTTGG 2160  
 Db 2213 ATGGGTACAGAGATTTACCAAGAAATTTGAAGTAACTGTAAGTATTTGATTCATCTTTGG 2272  
 Qy 2161 GTGGCTACTGTTTGTGTGATCTCTGAGACCTGGGCTTGAGACTAACACTACTCCAGTG 2220  
 Db 2273 GTGGCTACTGTTTGTGTGATCTCTGAGACCTGGGCTTGAGACTAACACTACTCCAGTG 2332  
 Qy 2221 GAGCTCCAGACCTTCCAAACATTTTTCATTTTTCATCTTCCCTACTCTGTATTCAGA 2280  
 Db 2333 GAGCTCCAGACCTTCCAAACATTTTTCATTTTTCATCTTCCCTACTCTGTATTCAGA 2392  
 Qy 2281 GGTGAAGAAATTTGCTTTGAGAAATTAATTAATTTCAATTTATTTGAAAGATGCCAGTGT 2340  
 Db 2393 GGTGAAGAAATTTGCTTTGAGAAATTAATTAATTTCAATTTATTTGAAAGATGCCAGTGT 2452  
 Qy 2341 AAGGTATATTTGAGAAAGTACCAAAATTTGATTTCTAATGACTTCAAGTAAATTAAT 2400  
 Db 2453 AAGGTATATTTGAGAAAGTACCAAAATTTGATTTCTAATGACTTCAATGAAATTAAT 2512  
 Qy 2401 GGCACAGGCGCACAGAGACCCCTTGTGTTCCAGTGAAGATGGGGCAACTGTTCTTTT 2460  
 Db 2513 GGCACAGGCGCACAGAGACCCCTTGTGTTCCAGTGAAGATGGGGCAACTGTTCTTTT 2572  
 Qy 2461 CCCATCAGGCGCAACATCTGGAGAAATTTCTTATCAGATCAGATCAGCTTTTCAACCACT 2520  
 Db 2573 CCCATCAGGCGCAACATCTGGAGAAATTTCTTATCAGATCAGATCAGCTTTTCAACCACT 2632  
 Qy 2521 GCTTCTGATGCTGTCAACCAAGATTTTATGTAAGCTGTAAGAAATAGAAATATCAAT 2580  
 Db 2633 GCTTCTGATGCTGTCAACCAAGATTTTATGTAAGCTGTAAGAAATATCAAT 2692  
 Qy 2581 TCACATCATCTTATTAAGTCTTGAATGATGCTAATAGGCTACAGATACCTGTAAGAACTTGG 2640  
 Db 2693 TCACATCATCTTATTAAGTCTTGAATGATGCTAATAGGCTACAGATACCTGTAAGAACTTGG 2752  
 Qy 2641 AGTTTCTATTTCTCTTAATACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2700  
 Db 2753 AGTTTCTATTTCTCTTAATACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2812  
 Qy 2701 GGAATGTTCTTGTCTCTTCAATCAATGAGCTTACCTCATTTGATTCGATTCCTTATGGC 2760  
 Db 2813 GGAATGTTCTTGTCTCTTCAATCAATGAGCTTACCTCATTTGATTCGATTCCTTATGGC 2872  
 Qy 2761 TGTGGTGAACAGAACATGATAATTTGCTCCAAATTTTACATTTTGAATTAATCTGACT 2820  
 Db 2873 TGTGGTGAACAGAACATGATAATTTGCTCCAAATTTTGAATTAATCTGACT 2932  
 Qy 2821 AAAAAGAAACAATGACAGATTAATTTGAAAGAAAGCTCTTTCATTTTGAAGCAAGGT 2880  
 Db 2933 AAAAAGAAACAATGACAGATTAATTTGAAAGAAAGCTCTTTCATTTTGAAGCAAGGT 2992  
 Qy 2881 TACCAAGAGAACTTCTATATCAAGAGGAAAGTGGCTTTTCAAGTCTTTTGGAAATTAAT 2940  
 Db 2993 TACCAAGAGAACTTCTATATCAAGAGGAAAGTGGCTTTTCAAGTCTTTTGGAAATTAAT 3052  
 Qy 2941 GACCCCTTGGAGAGCACTTGTGTGACGTTTGTTTTAAAGATGTTTCTTGAAGCCGAT 3000  
 Db 3053 GACCCCTTGGAGAGCACTTGTGTGACGTTTGTTTTAAAGATGTTTCTTGAAGCCGAT 3112  
 Qy 3001 CCTTACATAGATTTATGATGAGATGTTTATACAGAAACATACCTTGGCTTAAAGCAAT 3060  
 Db 3113 CCTTACATAGATTTATGATGAGATGTTTATACAGAAACATACCTTGGCTTAAAGCAAT 3172  
 Qy 3061 CAGAAATCCAAACGTTGAATTTTGGATCCAGAAAGATGATTCATAGTGAAGCTTCAAGGT 3120  
 Db 3173 CAGAAATCCAAACGTTGAATTTTGGATCCAGAAAGATGATTCATAGTGAAGCTTCAAGGT 3232

Qy 3121 GGCATTAAGATCCAGTAACACTTACAGCCTATATTTGTAATTTCTCTCGGATATAGA 3180  
 Db 3233 GGCATTAAGATCCAGTAACACTTACAGCCTATATTTGTAATTTCTCTCGGATATAGA 3292  
 Qy 3181 AAGTATCAG 3189  
 Db 3293 AAGTATCAG 3301  
 Db 3293 AAGTATCAG 3301  
 RESULT 15  
 ABQ79969  
 ID ABQ79969 standard; cDNA, 3535 BP.  
 XX  
 AC ABQ79969;  
 XX  
 DT 23-DEC-2002 (first entry)  
 XX  
 DE Human CD109 K15 variant protein encoding cDNA.  
 XX  
 KW CD109; alpha2 macroglobulin; thioester; cerebroprotective; cardiac;  
 KW immunosuppressive; haemostatic; anticoagulant; thrombolytic; human;  
 KW cardiovascular; vasotropic; gene therapy; CD109 K15; variant; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 113..4450  
 FT /tag= a  
 FT /product= "CD109 K15 variant"  
 PN WO200270696-A2.  
 PD 12-SEP-2002.  
 XX 07-MAR-2002; 2002WO-CA00292.  
 XX 07-MAR-2001; 2001US-273814P.  
 PR (SCHU/) SUTHERLAND R D.  
 PA (SUTR/) SUTHERLAND R D.  
 XX Schuh A, Sutherland RD;  
 PI MPI. 2002-713450/77.  
 DR P-PSDB; ABB82170.  
 DR  
 XX  
 PT New CD109 nucleic acids and polypeptides, useful in gene therapy,  
 PT particularly for treating strokes, myocardial infarctions, thrombosis,  
 PT thrombocytopenia, autoimmune diseases, or organ or bone marrow  
 PT transplantation -  
 XX  
 PX Claim 1; Fig 4b; 156bp; English.  
 XX  
 CC The invention relates to isolated nucleic acid molecules encoding CD109  
 CC polypeptides. These nucleic acid molecules include the human cDNA  
 CC sequences comprising CD109 K1, CD109 K1-H7, CD109 K15 or their variants.  
 CC CD109 is a novel member of the alpha2 macroglobulin (alpha2M)/C3, C4, C5  
 CC family of thioester-containing proteins. The CD109 polypeptides can be  
 CC expressed by standard recombinant methodology. The CD109 nucleic acid,  
 CC CD109 polypeptide or its mimetic, or the anti-CD109 antibody is useful  
 CC for treating or detecting a disease or disorder, e.g. conditions  
 CC associated with endothelial activation, platelet activation, activation  
 CC of the coagulation or fibrinolytic systems, activation of T lymphocytes  
 CC and of the complement system, quantitative or qualitative abnormalities  
 CC of platelet function, increased or impaired platelet aggregation and  
 CC activation, increased or impaired activation of the coagulation and/or  
 CC fibrinolytic systems, or impaired or increased immune activation. These  
 CC are also useful for treating cardiovascular disorders, stroke, myocardial  
 CC infarction, thrombosis, embolism, peripheral vascular disease,  
 CC thrombocytopenia, thrombocytopenia, autoimmune diseases, organ  
 CC transplantation, or bone marrow transplantation. The CD109 nucleic acids  
 CC or their antisense nucleotide sequence are useful in gene therapy for

treating these conditions. The present sequence represents the human  
CD109 K15 variant cDNA sequence.

Sequence 3535 BP; 1107 A; 676 C; 731 G; 1021 T; 0 other;

Query Match 74.3%; Score 3185.8; DB 24; Length 3535;

Basic Local Similarity 99.9%; Pred. No. 0;

Matches 3187; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Oy 1 ATGCGAGGCGCCAGCTCTGTGACGCGCGCCAGCTCTGTGAGTGTGACCGCGCGCTG 60
Db 113 ATGCGAGGCGCCAGCTCTGTGACGCGCGCCAGCTCTGTGAGTGTGACCGCGCGCTG 172
Oy 61 GCCGTGCTCCCGGGCTCGGTTTCTGTGTGACGCGCGCGAGATCATGAGCGCGAGGA 120
Db 173 GCCGTGCTCCCGGGCTCGGTTTCTGTGTGACGCGCGAGGATCATGAGCGCGAGGA 232
Oy 121 AATGTGACTATTGGGGGTGAGCTTGTGACACCTGCTTCACTGATCTGTGAGGCG 180
Db 223 AATGTGACTATTGGGGGTGAGCTTGTGACACCTGCTTCACTGATCTGTGAGGCG 292
Oy 181 GAGCTGCTCAAGACAGCATCAAACTCTGCTGTCTGTGAGAGAGAGAGTCTT 240
Db 293 GAGCTGCTCAAGACAGCATCAAACTCTGCTGTCTGTGAGAGAGAGTCTT 352
Oy 241 GAAAAAGGCTCTTTTAAGACACTTACTTTCATCTACTACTCTGACAGTGCAGATGAG 300
Db 353 GAAAAAGGCTCTTTTAAGACACTTACTTTCATCTACTACTCTGACAGTGCAGATGAG 412
Oy 301 ATTATGAGCTACGTGTACCGGAGCTACCGAGATGAGATTTTATTTCTTAATGATACC 360
Db 413 ATTATGAGCTACGTGTACCGGAGCTACCGAGATGAGATTTTATTTCTTAATGATACC 472
Oy 361 CGCTTATGCTTGAAGACCAAGAGATATCTGCTTCAATTCATAACAGACAGAGCTTATAC 420
Db 473 CGCTTATGCTTGAAGACCAAGAGATATCTGCTTCAATTCATAACAGAGAGCTTATAC 532
Oy 421 AAGCCAAAGCAAGAGTGAAGTTTGCATTTGTAACACTCTTCTGAGATTTTAAAGCTTAC 480
Db 533 AAGCCAAAGCAAGAGTGAAGTTTGCATTTGTAACACTCTTCTGAGATTTTAAAGCTTAC 592
Oy 481 AAAACCTCTTTAAACATTTCTCATTAAGAGACCCCAATTAATTTGATCAACAGTGTG 540
Db 593 AAAACCTCTTTAAACATTTCTCATTAAGAGACCCCAATTAATTTGATCAACAGTGTG 652
Oy 541 TCACAACAAGAGTCTTGAAGTCAATTCGAAACCTTTGAGTATCTTCCATCCATA 600
Db 653 TCACAACAAGTATCTTGAAGTCAATTCGAAACCTTTGAGTATCTTCCATCCATA 712
Oy 601 CTGTGTCAGTGTCTATCAAGTTCAAGTGAATGACAGACATTAATCAATCAATTCAG 660
Db 713 CTGTGTCAGTGTCTATCAAGTTCAAGTGAATGACAGACATTAATCAATCAATTCAG 772
Oy 661 GTTTCAGAAATATGATTAACCAAAATTTGAAGTGAAGTTCGACAGACCAATTAATGTTCT 720
Db 773 GTTTCAGAAATATGATTAACCAAAATTTGAAGTGAAGTTCGACAGACCAATTAATGTTCT 832
Oy 721 ATGATTTGAAGCAATTAATGATGATCAATCAAGAGATTAATGATGATGATGATGATG 780
Db 833 ATGATTTGAAGCAATTAATGATGATCAATCAAGAGATTAATGATGATGATGATGATG 892
Oy 781 AAGAGAGAGCTTACGCTTACATTTTACCTTATCTTTTGGGAAAGAAAGAAATATTT 840
Db 893 AAGAGAGAGCTTACGCTTACATTTTACCTTATCTTTTGGGAAAGAAAGAAATATTT 952
Oy 841 ACAAAAACATTTAAGATTAATGATGATGCAAACTTCTTTTAATGATGATGATGATGATG 900
Db 953 ACAAAAACATTTAAGATTAATGATGATGCAAACTTCTTTTAATGATGATGATGATGATG 1012
Oy 901 AATGTAATGATTTCTTGAATGATGATTTGATGATGATGATGATGATGATGATGATGATG 960
Db 1013 AATGTAATGATTTCTTGAATGATGATTTGATGATGATGATGATGATGATGATGATGATG 1072
```

```
Oy 961 GTAGAAATTTTAAACACAGTGAAGATTCAGTATCTTCAAGAAATGTAAGCACT 1020
Db 1073 GTAGAAATTTTAAACACAGTGAAGATTCAGTATCTTCAAGAAATGTAAGCACT 1132
Oy 1021 AATGTGCTTCAAGCAACATGATTCATGATGATTTTATGATTAATGATTAATGATTAATG 1080
Db 1133 AATGTGCTTCAAGCAACATGATTCATGATGATTTTATGATTAATGATTAATGATTAATG 1192
Oy 1081 AAGCATCTCTCAACTTCAAGCAGTGTGAGTGAAGTGAATCTGTGCTGAGCAACCACTG 1140
Db 1193 AAGCATCTCTCAACTTCAAGCAGTGTGAGTGAAGTGAATCTGTGCTGAGCAACCACTG 1252
Oy 1141 ACTTTGAAGAAAGAAATTAATGATGATTAATGATGATGATGATGATGATGATGATGATG 1200
Db 1253 ACTTTGAAGAAAGAAATTAATGATGATTAATGATGATGATGATGATGATGATGATGATG 1312
Oy 1201 TACTGAGCGGATCTTAACAGTGAATTCAGAAATGGAAGCTGTCAGAAATTAATTAAT 1260
Db 1313 TACTGAGCGGATCTTAACAGTGAATTCAGAAATGGAAGCTGTCAGAAATTAATTAAT 1372
Oy 1261 ACTGCCCCCAAGTGAAGCTTTTAAGATTAATGATTCATCTGAGAGATTCAGTGAAG 1320
Db 1373 ACTGCCCCCAAGTGAAGCTTTTAAGATTAATGATTCATCTGAGAGATTCAGTGAAG 1432
Oy 1321 CTACAGTTGAAGGCTTATTTCTTGTGATGAAGATGACATGACATGATCTGTTT 1380
Db 1433 CTACAGTTGAAGGCTTATTTCTTGTGATGAAGATGACATGACATGATCTGTTT 1492
Oy 1381 AAGTCTCTGATGAAGATTCATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 1440
Db 1493 AAGTCTCTGATGAAGATTCATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 1552
Oy 1441 TCGCCTTTTGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1500
Db 1553 TCGCCTTTTGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1612
Oy 1501 GATTCAGGAGGACAGTGTGCTGTGAGAAACAAATTCATCAATGTTCTCTTTAACA 1560
Db 1613 GATTCAGGAGGACAGTGTGCTGTGAGAAACAAATTCATCAATGTTCTCTTTAACA 1672
Oy 1561 CCAAGAAATCTTGTGACTCCAAAGCTGTGATTAATGATTAATGATTAATGATTAATGAT 1620
Db 1673 CCAAGAAATCTTGTGACTCCAAAGCTGTGATTAATGATTAATGATTAATGATTAATGAT 1732
Oy 1621 GAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1680
Db 1733 GAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1792
Oy 1681 CTATATTTGAAGTAAAGTGAAGTGAACCATCTGAGAAAGTCTCTTGAAGATCTGTTG 1740
Db 1793 CTATATTTGAAGTAAAGTGAAGTGAACCATCTGAGAAAGTCTCTTGAAGATCTGTTG 1852
Oy 1741 ACAAGGCTTACCTCATAGTTGGGATTTGATGCTGTGACAAAGTGTGAATCTGATGAAT 1800
Db 1853 ACAAGGCTTACCTCATAGTTGGGATTTGATGCTGTGACAAAGTGTGAATCTGATGAAT 1912
Oy 1801 GCTCTTAATGATTAATCAATGAAATGATGATGATGATGATGATGATGATGATGATGATG 1860
Db 1913 GCTCTTAATGATTAATCAATGAAATGATGATGATGATGATGATGATGATGATGATGATG 1972
Oy 1861 TATTAATTTAAGGATGTCATGAATTTCTTGTGAGTCTTCAAGAAATGAGATCTGAGTA 1920
Db 1973 TATTAATTTAAGGATGTCATGAATTTCTTGTGAGTCTTCAAGAAATGAGATCTGAGTA 2032
Oy 1921 TTGACAGATGCAAACTCAAGAGATTAATTAATGATGATGATGATGATGATGATGATGAT 1980
Db 2033 TTGACAGATGCAAACTCAAGAGATTAATTAATGATGATGATGATGATGATGATGATGAT 2092
Oy 1981 GCTGAGAGGTTTATGAGAGAAATGAAAGACATATGATGATGATGATGATGATGATGATG 2040
Db 2093 GCTGAGAGGTTTATGAGAGAAATGAAAGACATATGATGATGATGATGATGATGATGATG 2152
Oy 2041 GGTAGAGTCCACATGTCGAAAGCATTTTCCAGAGACTTGAATTTGGCTTGAACCCMAC 2100
```

Db 2153 GGTAGAGTCCATGTCGGAAGCATTTTCCAGACCTTGATGGCTAGACACAAAC 2212  
Qy 2101 ATGGGTTACAGGATTTACCAAGAAATTGGAATACGTACCTGATCTTCACTTCTGG 2160  
Db 2213 ATGGGTTCCAGATTTTACCAAGAAATTGGAATACGTACCTGATCTTCACTTCTGG 2272  
Qy 2161 GTGGCTACTGGTTTGTGATCTCTGAGACCTGGGCTTTGGAATCAACTACTCCAGTG 2220  
Db 2273 GTGGCTACTGGTTTGTGATCTCTGAGACCTGGGCTTTGGAATCAACTACTCCAGTG 2332  
Qy 2221 GAGCTCCAGCCTTCCAACTTTTTCATTTTGTGAATCTTCCCTACTCTGTATCAGA 2280  
Db 2333 GAGCTCCAGCCTTCCAACTTTTTCATTTTGTGAATCTTCCCTACTCTGTATCAGA 2392  
Qy 2281 GGTGAAGAAATTTGCTTGGAAATACATATTAATTAATTTGAAAGATGCCACTGAGTT 2340  
Db 2393 GGTGAAGAAATTTGCTTGGAAATACATATTAATTAATTTGAAAGATGCCACTGAGTT 2452  
Qy 2341 AAGGTAAATCATGAGAAAGTGACAAATTTGATATCTAATGACTTCAAGTGAATTAAT 2400  
Db 2453 AAGGTAAATCATGAGAAAGTGACAAATTTGATATCTAATGACTTCAAGTGAATTAAT 2512  
Qy 2401 GCCACAGGCCACACAGACACCTTCTGTGCCAGTGAAGATGGGCAACTGTTCTTTT 2460  
Db 2513 GCCACAGGCCACACAGACACCTTCTGTGCCAGTGAAGATGGGCAACTGTTCTTTT 2572  
Qy 2461 CCCATAGGCGCAACATCTGGGAGAAATTCCTATCAAGTCAAGCTCTTTCACCCACT 2520  
Db 2573 CCCATAGGCGCAACATCTGGGAGAAATTCCTATCAAGTCAAGCTCTTTCACCCACT 2632  
Qy 2521 GCTTCTGATGCTGTCAACCCAGATGATTTTAAAGGCTGAAGGAATAGAAAAATCATAT 2580  
Db 2633 GCTTCTGATGCTGTCAACCCAGATGATTTTAAAGGCTGAAGGAATAGAAAAATCATAT 2692  
Qy 2581 TCACATTCATCTTATTTAAGCTTGAATGACATAGGCTACAGATACCTGAAACTTTG 2640  
Db 2693 TCACATTCATCTTATTTAAGCTTGAATGACATAGGCTACAGATACCTGAAACTTTG 2752  
Qy 2641 AGTTTCTATTCCTCTTAATACAGTACGTCGGAGTGAAGAAAGTTGAGATCACTGCAATT 2700  
Db 2753 AGTTTCTATTCCTCTTAATACAGTACGTCGGAGTGAAGAAAGTTGAGATCACTGCAATT 2812  
Qy 2701 GAGATGTTCTTGTGCTTCCATCAATGCTTAGCTCATTTGATGGATGCTTATGGC 2760  
Db 2813 GAGATGTTCTTGTGCTTCCATCAATGCTTAGCTCATTTGATGGATGCTTATGGC 2872  
Qy 2761 TGTGTGAACAGAACATGATTAATTTTGTCCAAATATTTTACATTTTGAATTAATCTGACT 2820  
Db 2873 TGTGTGAACAGAACATGATTAATTTTGTCCAAATATTTTACATTTTGAATTAATCTGACT 2932  
Qy 2821 AAAAAAGAAACACTGACAGATTAATTTGAAAGAAAAAGCTTTTCATTTATGAGGCAAGGT 2880  
Db 2933 AAAAAAGAAACACTGACAGATTAATTTGAAAGAAAAAGCTTTTCATTTATGAGGCAAGGT 2992  
Qy 2881 TACACAGAGAACTTCTATACAGAGGAAGATGCTCTTTCAGTCTTTTGGGAATTAAT 2940  
Db 2993 TACACAGAGAACTTCTATACAGAGGAAGATGCTCTTTCAGTCTTTTGGGAATTAAT 3052  
Qy 2941 GACCTTCTGGGAGCATTTGTTGTGACGTTTGTGTTAAGATGTTTCTTGAAGCCGAT 3000  
Db 3053 GACCTTCTGGGAGCATTTGTTGTGACGTTTGTGTTAAGATGTTTCTTGAAGCCGAT 3112  
Qy 3001 CCTTACATAGATTTTATGATGAATGTGTTACACAGAACATACACTTGCTTAAAGGACAT 3060  
Db 3113 CCTTACATAGATTTTATGATGAATGTGTTACACAGAACATACACTTGCTTAAAGGACAT 3172  
Qy 3061 CAGAAATCCAAAGGTAATTTTGGATCCAGGAAGATGATTCATAGTGAAGCTTCAAGGT 3120  
Db 3173 CAGAAATCCAAAGGTAATTTTGGATCCAGGAAGATGATTCATAGTGAAGCTTCAAGGT 3232  
Qy 3121 GGCATATAAAGTCCAGTAAACATTACAGCTATATTTGTAATTTCTCTCTGGGATATAGA 3180

Db 3233 GGCATATAAAGTCCAGTAACTTACAGCTATATTTGTAATTTCTCTCTGGGATATAGA 3292  
Qy 3181 AAGTATCAG 3189  
Db 3293 AAGTATCAG 3301

Search completed: January 16, 2004, 14:38:44  
Job time : 1046 secs

**THIS PAGE BLANK (08970)**

Human platelet alpha-2-macroglobulin
Hydrophobic domain
Human blood cell s.
Human r150 DNA #2.1
Human CDI09 R15 part
Human CDI09 R15 part
Human polynucleotide
Human secreted protein
Alpha-1-proteinase inhibitor
Human proctase and
Human prostate acid
Angiogenesis-related
Human gene expression
DNA encoding novel
Human gene expression
Bovine EST associated
Human gene expression
C. elegans alpha-2-macroglobulin
Drosophila melanogaster
Drosophila melanogaster
Drosophila melanogaster
Drosophila melanogaster
Drosophila melanogaster
Drosophila melanogaster
DNA encoding novel
Gene #3730 used to clone
DNA encoding novel
Human ORF3 p.f.1
DNA encoding novel
Human alpha-2-macroglobulin
Human pregnancy zell
Human platelet alpha-2-macroglobulin
DNA encoding human
Human cDNA SEQ ID
Human polynucleotide

**THIS PAGE BLANK (USPTO)**

XX (UNMC-) UNIV MCGILL.  
PA Philip A, Tam B;  
PI WPI, 2003-093100/08.  
XX P-PSDB; AAB32012.

DR Novel transforming growth factor (TGF)-beta 1 binding reagent which  
XX comprises r150 protein which acts as accessory receptor of TGF-beta,  
PT useful for negatively modulating TGF-beta activity, and thus for  
PT treating cancer

PS Claim 7, Page 91-93; 127p; English.

XX The invention relates to novel transforming growth factor (TGF)-beta 1  
XX binding reagent which comprises a glycosylphosphatidyl inositol (GPI)-  
XX anchored TGF-beta binding protein referred to as r150 which acts as  
XX an accessory receptor of TGF-beta. The invention is used for negatively  
XX modulating TGF-beta activity, and thus for treating conditions  
XX characterised by overproduction of TGF-beta, such as cancer. Antisense  
XX molecules of the invention are used for increasing TGF-beta availability  
XX and increase graft success. The present sequence is a gene encoding  
XX human r150 protein.

XX Sequence 4369 BP; 1301 A; 842 C; 932 G; 1293 T; 1 other;

Query Match 99.9%; Score 4281.2; DB 25; Length 4369;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 4283; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGACAGGCGCCACCGCTCTGACCGCGCCACCTCTGCGGTGACCGCGCGCTG 60  
DB 26 ATGACAGGCGCCACCGCTCTGACCGCGCCACCTCTCTGCGGTGACCGCGCGCTG 85  
QY 61 GCGGTGCGCTCCGCGCGCTGCGGTGCGGTGACCGCGCGCGCGCGCGCGCGCG 120  
DB 86 GCGGTGCGCTCCGCGCGCTGCGGTGCGGTGACCGCGCGCGCGCGCGCGCGCG 145  
QY 121 AATGCTACTATTGGGGGTGAGCTTCTGAAACCTGCGCTTCAAGGTGACTGAAAGCG 180  
DB 146 AATGCTACTATTGGGGGTGAGCTTCTGAAACCTGCGCTTCAAGGTGACTGAAAGCG 205  
QY 181 GAGTCTCTCAAGACAGCATCAAACTCACTGCTCTGCTCTGAAAGCAGAGAGCTTT 240  
DB 206 GAGTCTCTCAAGACAGCATCAAACTCACTGCTCTGCTCTGAAAGCAGAGAGCTTT 265  
QY 241 GAAAGAGCTCTTTTAAAGACACTTACTCTTCCATCACTGCTGAAAGCAGAGAG 300  
DB 266 GAAAGAGCTCTTTTAAAGACACTTACTCTTCCATCACTGCTGAAAGCAGAGAG 325  
QY 301 AATTAGAGCTAGCTGTAACCGGAGTACCGAGAGAGAGATTATCTTAATAGTACC 360  
DB 326 AATTAGAGCTAGCTGTAACCGGAGTACCGAGAGAGAGATTATCTTAATAGTACC 385  
QY 361 CGCTTACTTTGAGACCAAGAGATATCTGCTTCAATCAAGACAGAGCGCTTATAC 420  
DB 386 CGCTTACTTTGAGACCAAGAGATATCTGCTTCAATCAAGACAGAGCGCTTATAC 445  
QY 421 AAGCCAAAGCAG 480  
DB 446 AAGCCAAAGCAG 505  
QY 481 AAGCCAAAGCAG 540  
DB 506 AAGCCAAAGCAG 565  
QY 541 TCAACAACAAGAGATCTGAGAGATCTTCAAACTTTTCAAGTATCTTCCATCAATA 600  
DB 566 TCAACAACAAGAGATCTTGAAGATCTTCAAACTTTTCAAGTATCTTCCATCAATA 625  
QY 601 CTGAGAGAGAGAGATCTTCAAGTATCTTCAAGTATCTTCCATCAATA 660

DB 626 CTGAGAGAGAGAGATCTTCAAGTATCTTCAAGTATCTTCCATCAATA 685  
QY 661 GTTTCAGAAATATGATTAACCAAAATTTGAAGTACTTTGAGACACACCATTAATTTCT 720  
DB 686 GTTTCAGAAATATGATTAACCAAAATTTGAAGTACTTTGAGACACACCATTAATTTCT 745  
QY 721 ATGAATTTAAGCATTTAATGATGACATCAAGGCAAGATATATATGGAAGCCAGTG 780  
DB 746 ATGAATTTAAGCATTTAATGATGACATCAAGGCAAGATATATATGGAAGCCAGTG 805  
QY 781 AAAG 840  
DB 806 AAAG 865  
QY 841 ACAAATTTAAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
DB 866 ACAAATTTAAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 925  
QY 901 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
DB 926 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 985  
QY 961 GTAGAAATTTTAAACACAG 1020  
DB 986 GTAGAAATTTTAAACACAG 1045  
QY 1021 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080  
DB 1046 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1105  
QY 1081 AAGCATGCTCAACTTCAACAGCAGCAGTGTGAGTAACTGTGTGATGAGACCAACTG 1140  
DB 1106 AAGCATGCTCAACTTCAACAGCAGCAGTGTGAGTAACTGTGTGATGAGACCAACTG 1165  
QY 1141 ACTCTGAAG 1200  
DB 1166 ACTCTGAAG 1225  
QY 1201 TACTGAGCGAGATTAACAGTGAATCAGAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260  
DB 1226 TACTGAGCGAGATTAACAGTGAATCAGAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1285  
QY 1261 ACTGCTCCCAAG 1320  
DB 1286 ACTGCTCCCAAG 1345  
QY 1321 CTACAGTGAAG 1380  
DB 1346 CTACAGTGAAG 1405  
QY 1381 AAGTCTCTTGAAGACATATCACTCAATCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440  
DB 1406 AAGTCTCTTGAAGACATATCACTCAATCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1465  
QY 1441 TCGCTTTTGAAGT 1500  
DB 1466 TCGCTTTTGAAGT 1525  
QY 1501 GTATTCAG 1560  
DB 1526 GTATTCAG 1585  
QY 1561 CCAGAAATTTCTTGAAGT 1620  
DB 1586 CCAGAAATTTCTTGAAGT 1645  
QY 1621 GAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680  
DB 1646 GAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1705  
QY 1681 CTATTTGAG 1740  
DB 1706 CTATTTGAG 1765



**THIS PAGE BLANK (USPTO)**

1741 ACACAGCCTGACCTCAGATGTTGGAGTTTGAAGCTTTGACAAAGTGTGATCTGATGAT 1800  
1766 ACACAGCCTGACCTCAGATGTTGGAGTTTGAAGCTTTGACAAAGTGTGATCTGATGAT 1825  
1801 GCTCTTAATGATATTAATGAGAAATGATGTCATGATGATGAACTTATATACACAGA 1860  
1826 GCTCTTAATGATATTAATGAGAAATGATGTCATGATGATGAACTTATATACACAGA 1885  
1861 TATTATTTAGGCACTTCAATGAAATCTTTTGAAGTCTTTCAGAAATGTGAACTTGGGTA 1920  
1886 TATTATTTAGGCACTTCAATGAAATCTTTTGAAGTCTTTCAGAAATGTGAACTTGGGTA 1945  
1921 TTGAAGATGCAACCTCAGAGAGATTTATTTAGATGAGTTTATGACAAATGCAAAAT 1980  
1946 TTGAAGATGCAACCTCAGAGAGATTTATTTAGATGAGTTTATGACAAATGCAAAAT 2005  
1981 GCTGAGAGTTTATGAGAGAAATGAGAGACATATTTAGATATTTGATGATCTTTCTTG 2040  
2006 GCTGAGAGTTTATGAGAGAAATGAGAGACATATTTAGATATTTGATGATCTTTCTTG 2065  
2041 GGTAGAGTCCACATGTCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2100  
2066 GGTAGAGTCCACATGTCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2125  
2101 ATGGGTTACAGATTTTACAGAAATTTGAGATGATCTGATCTTATCTACTCTTG 2160  
2126 ATGGGTTACAGATTTTACAGAAATTTGAGATGATCTGATCTTATCTACTCTTG 2185  
2161 GTGGCTACTGTTTGTGATCTCTGAGAGCTGGGCTTGAAGTAACTCACTCAAGT 2220  
2186 GTGGCTACTGTTTGTGATCTCTGAGAGCTGGGCTTGAAGTAACTCACTCAAGT 2245  
2221 GAGTCCAAAGCTTCCAAACATTTTCAATTTTGAATCTTCCACTCTGATACAGA 2280  
2246 GAGTCCAAAGCTTCCAAACATTTTCAATTTTGAATCTTCCACTCTGATACAGA 2305  
2281 GGTGAGAAATTTGCTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2340  
2306 GGTGAGAAATTTGCTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2365  
2341 AAGGTAATCATTAAGAAAGTGAACAAATTTGATTTCTAATGATCTCAAGTAAATTAAT 2400  
2366 AAGGTAATCATTAAGAAAGTGAACAAATTTGATTTCTAATGATCTCAAGTAAATTAAT 2425  
2401 GCCACAGGCAACAGAGACCTTCTGATGTCAGAGAGATGGGCACTGTTCTTTT 2460  
2426 GCCACAGGCAACAGAGACCTTCTGATGTCAGAGAGATGGGCACTGTTCTTTT 2485  
2461 CCCATCAGGCAACAGATCTGGGAAATTTCTATCAAGTCAAGCTCTTCAACCACT 2520  
2486 CCCATCAGGCAACAGATCTGGGAAATTTCTATCAAGTCAAGCTCTTCAACCACT 2545  
2521 GCTTCTGATGCTGTCAACCAAGATGATTTTATTAAGGCTGAAGAAATGAATAATCAT 2580  
2546 GCTTCTGATGCTGTCAACCAAGATGATTTTATTAAGGCTGAAGAAATGAATAATCAT 2605  
2581 TCACATTCATCTTATTAAGCTTGAAGTCAATAGGCTCAAGAGTCAAGCTCTTCAACCACT 2640  
2606 TCACATTCATCTTATTAAGCTTGAAGTCAATAGGCTCAAGAGTCAAGCTCTTCAACCACT 2665  
2641 AGTTTCTCATTTCTCTTAATGATGATGATGATGATGATGATGATGATGATGATGAT 2700  
2666 AGTTTCTCATTTCTCTTAATGATGATGATGATGATGATGATGATGATGATGATGAT 2725  
2701 GAGATGCTCTGCTGCTTCCATCAATGCTTGAAGCTTGAAGTGAAGTGAAGTGAAGTGAAG 2760  
2726 GAGATGCTCTGCTGCTTCCATCAATGCTTGAAGCTTGAAGTGAAGTGAAGTGAAGTGAAG 2785  
2761 TGTGTTGAACAGAAATGATTAATTTTGTCTCAATTAATTTTGAATTTTGAATTTTGAAT 2820  
2786 TGTGTTGAACAGAAATGATTAATTTTGTCTCAATTAATTTTGAATTTTGAATTTTGAAT 2845

2821 AAAAGAAACAACTGACAGATTAATTTGAAAGAAAGCTCTTCAATTTATGAGCAAGT 2880  
2846 AAAAGAAACAACTGACAGATTAATTTGAAAGAAAGCTCTTCAATTTATGAGCAAGT 2905  
2881 TACAGAGAAAGCTCTCTATCAGAGGAAAGATGCTCTTCAAGTCTTTGGGAAATTAAT 2940  
2906 TACAGAGAAAGCTCTCTATCAGAGGAAAGATGCTCTTCAAGTCTTTGGGAAATTAAT 2965  
2941 GACCTTCTGGAGCACTTGGTGTACGCTTTTGTATTAAGATGTTCTTGAAGCCAT 3000  
2966 GACCTTCTGGAGCACTTGGTGTACGCTTTTGTATTAAGATGTTCTTGAAGCCAT 3025  
3001 CCTTACATTAATTTGATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 3060  
3026 CCTTACATTAATTTGATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 3085  
3061 CAGAAATCCAGAGGTAATTTGGGATCCAGAGAGATGATTAAGTGAAGCTTCAAGT 3120  
3086 CAGAAATCCAGAGGTAATTTGGGATCCAGAGAGATGATTAAGTGAAGCTTCAAGT 3145  
3121 GGCATTAAGTCCAGATTAACCTTACAGCTTATATTTGAATCTTCTCTGGATATAGA 3180  
3146 GGCATTAAGTCCAGATTAACCTTACAGCTTATATTTGAATCTTCTCTGGATATAGA 3205  
3181 AAGTATCAGCTTAACATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3240  
3206 AAGTATCAGCTTAACATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3265  
3241 AGAGAAATTTGACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3300  
3266 AGAGAAATTTGACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3325  
3301 AGTCTTAAGGAGAGAGAGCTTGAATTAATGATGATGATGATGATGATGATGATGATGAT 3360  
3326 AGTCTTAAGGAGAGAGAGCTTGAATTAATGATGATGATGATGATGATGATGATGATGAT 3385  
3361 ATGCAATCTGGGTCATCAAGATCCAACTTTTCTGATCTTGGAGCCAGCTCTCTG 3420  
3386 ATGCAATCTGGGTCATCAAGATCCAACTTTTCTGATCTTGGAGCCAGCTCTCTG 3445  
3421 GATATGAGTGAAGTGAAGCTTGAAGCTGCTCAAGCTTCAATTAATTAATTAATTAATTAAT 3480  
3446 GATATGAGTGAAGTGAAGCTTGAAGCTGCTCAAGCTTCAATTAATTAATTAATTAATTAAT 3505  
3481 GGAATCCCAATTTATGAGGTCCTAAGAGAGCAAGAAATAGCTTGGGTTTGGATCT 3540  
3506 GGAATCCCAATTTATGAGGTCCTAAGAGAGCAAGAAATAGCTTGGGTTTGGATCT 3565  
3541 ACTCAGATACCACTGTCCTTTAAAGCTCTGCTGAATTTGAGAGCTTGAATTAATTAATTAAT 3600  
3566 ACTCAGATACCACTGTCCTTTAAAGCTCTGCTGAATTTGAGAGCTTGAATTAATTAATTAAT 3625  
3601 GAAAGGCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3660  
3626 GAAAGGCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3685  
3661 GTAAGGCCAATGAGAGTAAATTTCCGCAAAATGATTTTGAATTTGATTTGATGATGAT 3720  
3686 GTAAGGCCAATGAGAGTAAATTTCCGCAAAATGATTTTGAATTTGATTTGATGATGAT 3745  
3721 AATGTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3780  
3746 AATGTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3805  
3781 CAAGAGCTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3840  
3806 CAAGAGCTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3865  
3841 GATTTGAATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3900  
3866 GATTTGAATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3925  
3901 GTTAACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3960

**THIS PAGE BLANK (USPTO)**

Db 3926 GTTAACTTATGAGCTTATGCTGCTGAGAACTTCTCTGAGCAGACAGT 3985  
Oy 3961 AAGAAAGTGAATGATGATGAGAAACCTCAACCTTATTTGATTTGTAATGAACC 4020  
Db 3986 AAGAAAGTGAATGATGATGAGAAACCTCAACCTTATTTGATTTGTAATGAACC 4045  
Oy 4021 CAGTTTGTGTAATATTCCTGCTGAGAACTTAAAGTTTCAATACCAAGATGCT 4080  
Db 4046 CAGTTTGTGTAATATTCCTGCTGAGAACTTAAAGTTTCAATACCAAGATGCT 4105  
Oy 4081 TCAAGTTCATAGTGAATTAATGAGCCAGAGAGACAGCGGTGAGAGATTAACCTT 4140  
Db 4106 TCAAGTTCATAGTGAATTAATGAGCCAGAGAGACAGCGGTGAGAGATTAACCTT 4165  
Oy 4141 GAAGTGAACCTGCTCCCTGCTGAGACCTTTGACATGATGTCAGAGGCTCTTGTGAG 4200  
Db 4166 GAAGTGAACCTGCTCCCTGCTGAGACCTTTGACATGATGTCAGAGGCTCTTGTGAG 4225  
Oy 4201 GATGAGCTTCAGAGCTCCCATCATCATCTTCAGTCACTTATTTTCTGTTCAAGCTT 4260  
Db 4226 GATGAGCTTCAGAGCTCCCATCATCATCTTCAGTCACTTATTTTCTGTTCAAGCTT 4285  
Oy 4261 CTGACTTTATGGAAGCTTTGCTGTGA 4287  
Db 4286 CTGACTTTATGGAAGCTTTGCTGTGA 4312

RESULT 2  
AB079964  
ID AB079964 standard; cDNA; 4761 BP.

XX AB079964;

DT 23-DEC-2002 (first entry)

XX Human CD109 K1 protein encoding cDNA.

XX CD109; alpha2 macroglobulin; thioester; cerebroprotective; cardiant;  
XX immunosuppressive; haemostatic; anticoagulant; thrombolytic; human;  
XX cardiovascular; vasotropic; gene therapy; CD109 K1; gene; ss.

XX Homo sapiens.

XX Key location/Qualifiers  
FH 113..4450  
FT /tag= a  
FT /product= "CD109 K1"

XX WO200270696-A2.

XX 12-SEP-2002.

XX 07-MAR-2002; 2002WO-CA00292.

XX 07-MAR-2001; 2001US-273814P.

XX (SCHU/) SCHUH A.  
XX (SUTH/) SUTHERLAND R D.

XX Schuh A, Sutherland RD;

XX MPI; 2002-713450/77.

XX P-PSDB; ABB82165.

XX New CD109 nucleic acids and polypeptides, useful in gene therapy,  
XX particularly for treating strokes, myocardial infarctions, thrombosis,  
XX thrombocytopenia, autoimmune diseases, or organ or bone marrow  
XX transplantation -

XX Claim 1; Fig 1a; 156pp; English.

XX The invention relates to isolated nucleic acid molecules encoding CD109

CC polypeptides. These nucleic acid molecules include the human cDNA  
CC sequences comprising CD109 K1, CD109 K1-H7, CD109 K1S or their variants.  
CC CD109 is a novel member of the alpha2 macroglobulin (alpha2m)/C3, C4, C5  
CC family of thioester-containing proteins. The CD109 polypeptides can be  
CC expressed by standard recombinant methodology. The CD109 nucleic acid,  
CC CD109 polypeptide or its mimetic, or the anti-CD109 antibody is useful  
CC for treating or detecting a disease or disorder, e.g. conditions  
CC associated with endothelial activation, platelet activation, activation  
CC of the coagulation or fibrinolytic systems, activation of T lymphocytes  
CC and of the complement system, quantitative or qualitative abnormalities  
CC of platelet function, increased or impaired platelet aggregation and  
CC activation, increased or impaired activation of the coagulation and/or  
CC fibrinolytic systems, or impaired or increased immune activation. These  
CC are also useful for treating cardiovascular disorders, stroke, myocardial  
CC infarction, thrombosis, embolism, peripheral vascular disease,  
CC thrombocytopenia, thrombocytopenia, autoimmune diseases, organ  
CC transplantation, or bone marrow transplantation. The CD109 nucleic acids  
CC or their antisense nucleic acid sequence are useful in gene therapy for  
CC treating these conditions. The present sequence represents the human  
CC CD109 K1 cDNA sequence.

SO Sequence 4761 BP; 1420 A; 912 C; 1016 G; 1413 T; 0 other;

Query Match 98.5%; Score 4224.4; DB 24; Length 4761;

Best Local Similarity 98.8%; Pred. No. 0;

Matches 4286; Conservative 0; Mismatches 1; Indels 51; Gaps 1;

Oy 1 ATGCAAGGCGCCACGCTCTGACCGCCACCTCCTGCTGCTGAGCAGCGCGCTG 60  
Db 113 ATGCAAGGCGCCACGCTCTGACCGCCACCTCCTGCTGCTGAGCAGCGCGCTG 172  
Oy 61 GCCGTGCTCCCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120  
Db 173 GCCGTGCTCCCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 232  
Oy 121 AATGTACTATTTGGGGTGGAGCTTCTGGAACATCGCCCTTCAAGGAGTGAAGGCG 180  
Db 233 AATGTACTATTTGGGGTGGAGCTTCTGGAACATCGCCCTTCAAGGAGTGAAGGCG 292  
Oy 181 GAGCTGCTCAAGACAGCATCAAACTCATCTGCTGCTGCTGCTGCTGCTGCTGCT 240  
Db 293 GAGCTGCTCAAGACAGCATCAAACTCATCTGCTGCTGCTGCTGCTGCTGCTGCT 352  
Oy 241 GAAAAAGCTCTTTTAAAGACACTTCTTCTCACTACTCTGAAACAGTCAATGAG 300  
Db 353 GAAAAAGCTCTTTTAAAGACACTTCTTCTCACTACTCTGAAACAGTCAATGAG 412  
Oy 301 ATTATGAGCTACGCTGTAACCGGACGTACCGAGATGAGATTTATCTTAATAGTACC 360  
Db 413 ATTATGAGCTACGCTGTAACCGGACGTACCGAGATGAGATTTATCTTAATAGTACC 472  
Oy 413 CGCTTATCATTTGAGACCAAGAGATATGCTTCAATTAACAGCAAGGCTTATAC 420  
Db 473 CGCTTATCATTTGAGACCAAGAGATATGCTTCAATTAACAGCAAGGCTTATAC 532  
Oy 421 AAGCCAAAGCAAGAGTGAAGTTTCCATTTGTAACCTCTTCAATTTAAAGCTTAC 480  
Db 473 CGCTTATCATTTGAGACCAAGAGATATGCTTCAATTAACAGCAAGGCTTATAC 532  
Oy 481 AAAACCTTTTAAACATTTCAATTAAGAGCCCAAAATCAATTTGATCCAAAGTGTG 540  
Db 533 AAAACCTTTTAAACATTTCAATTAAGAGCCCAAAATCAATTTGATCCAAAGTGTG 652  
Oy 541 TCAACAAAGATGATTTGAGAGTCTTCCAAATCTTCAAGTATCTTCCATCCATA 600  
Db 653 TCAACAAAGATGATTTGAGAGTCTTCCAAATCTTCAAGTATCTTCCATCCATA 712  
Oy 601 CTGCTGATCTGCTATTTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 660  
Db 713 CTGCTGATCTGCTATTTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 772  
Oy 661 GTTTCAGATATGATATCAAAATTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 720

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

nucleic search, using sw model

January 16, 2004, 12:38:20 ; Search time 14986 Seconds

(without alignments)  
11702.894 Million cell updates/sec

US-10-020-095-3

4287

Sequence: 1 atcgagggccacccgcctcct.....ttatggaacttggtgtgtga 4287

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

GenEmbl: \*  
1: gb\_ba: \*  
2: gb\_htg: \*  
3: gb\_in: \*  
4: gb\_om: \*  
5: gb\_ov: \*  
6: gb\_pat: \*  
7: gb\_ph: \*  
8: gb\_pl: \*  
9: gb\_pr: \*  
10: gb\_ro: \*  
11: gb\_str: \*  
12: gb\_sy: \*  
13: gb\_un: \*  
14: gb\_vl: \*  
15: gb\_ba: \*  
16: em\_fun: \*  
17: em\_hum: \*  
18: em\_in: \*  
19: em\_mu: \*  
20: em\_om: \*  
21: em\_or: \*  
22: em\_ov: \*  
23: em\_pat: \*  
24: em\_ph: \*  
25: em\_pl: \*  
26: em\_ro: \*  
27: em\_str: \*  
28: em\_un: \*  
29: em\_vl: \*  
30: em\_htg\_hum: \*  
31: em\_htg\_inv: \*  
32: em\_htg\_other: \*  
33: em\_htg\_mus: \*  
34: em\_htg\_pln: \*  
35: em\_htg\_rtd: \*  
36: em\_htg\_mam: \*  
37: em\_htg\_vit: \*  
38: em\_sy: \*  
39: em\_htg\_hum: \*  
40: em\_htg\_mus: \*  
41: em\_htg\_other: \*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	4224.4	98.5	4761	6	AX534934
2	4224.4	98.5	5883	9	AF410459
3	4224.4	98.5	5895	6	AX534938
4	4222.8	98.5	4761	6	AX534936
5	4222.8	98.5	5895	6	AX534940
6	4221.4	98.5	4335	6	AX537610
7	4219.8	98.4	4335	6	AX537612
8	3187.4	74.4	3535	6	AX534942
9	3185.8	74.3	3535	6	AX534944
10	2709.2	63.2	4735	10	AY083458
11	2709.2	63.2	5644	10	BC052443
12	2448.2	57.1	2938	9	HSW805581
13	1969.4	45.9	2273	9	AK095888
14	1969.4	45.9	1700	6	AX083146
15	232	5.4	163577	9	AL590428
16	230	5.1	5603	3	AC012408
17	219.6	4.8	4678	3	CIN431688
18	203.8	4.8	4678	3	D83196
19	182	4.2	251250	2	AC097023
20	177.4	4.1	149597	2	AC034271
21	175	4.1	91419	9	AL591480
22	174	4.1	4527	6	AX054953
23	174	4.1	4560	6	AX054954
24	171.6	4.0	2629	3	AY118302
25	167.6	3.9	4715	5	GGOVOM
26	167	3.9	4578	3	AY122084
27	167	3.9	4579	3	AY122084
28	160	3.7	144589	2	DME269539
29	160	3.7	191258	2	AC112668
30	155.4	3.6	4620	10	RN4113
31	154.8	3.6	176144	2	AC012408
32	154.2	3.6	4636	10	RAT4113A
33	152.2	3.6	4220	3	DME269538
34	151.2	3.5	4575	10	GPMSPB
35	149.8	3.5	2501	10	RAT4113
36	146.4	3.4	4950	5	XELENDI
37	143.8	3.4	122013	3	AC092399
38	143.8	3.4	164414	3	AC092395
39	143.8	3.4	168503	3	AC115484
40	143.8	3.4	275699	3	AE003649
41	143.8	3.4	337178	3	DROSADH08
42	140.8	3.3	191258	2	AC118254
43	136	3.2	4623	10	BC036299
44	136	3.2	4672	10	BC051037
45	134.8	3.1	2689	5	AB026130

## ALIGNMENTS

RESULT 1	AX534934	4761 bp	DNA	linear	PAT 22-NOV-2002
LOCUS	AX534934				
DEFINITION	Sequence 1 from Patent WO02070696.				
ACCESSION	AX534934				
VERSION	AX534934.1	GI:25261474			
KEYWORDS					
SOURCE					
ORGANISM	Homo sapiens (human)				
REFERENCE	Schuh, A. and Sutherland, R. D.				
AUTHORS	Cd109 nucleic acid molecules polypeptides and methods of use				
TITLE	Patent: WO 02070696-A 1 12-SEP-2002;				
JOURNAL					

**THIS PAGE BLANK (USPTO)**

Y	3670	ATGCGATTAATATTTCCGGAATGTTTGAATTTGCTATTTGCTCAATGTTGA	3729
b	3833	ATGCGATTAATATTTCCGGAATGTTTGAATTTGCTATTTGCTCAATGTTGA	3892
Y	3730	TATATATGAGGCTTCTGGGCTCTTCTAGAGAGCAAGATCTATCCAAATCAAGAGCC	3789
b	3893	TATATATGAGGCTTCTGGGCTCTTCTAGAGAGCAAGATCTATCCAAATCAAGAGCC	3952
Y	3790	TTTGATTTAGATGTTGCTGTAAAGAAATTAAGATGATCTCAATCATGTGATTTGAAT	3849
b	3953	TTTGATTTAGATGTTGCTGTAAAGAAATTAAGATGATCTCAATCATGTGATTTGAAT	4012
Y	3850	GTCGTGACAGCTTTTGGGCGCGGGTGAAGAGGCAATGCTTATAGAAATTAACCTTA	3909
b	4013	GTCGTGACAGCTTTTGGGCGCGGGTGAAGAGGCAATGCTTATAGAAATTAACCTTA	4072
Y	3910	TTAAGTGGCTTTATGAGTCTCTTCAAGAACATTTCTCTGACGAGACAGTGAAGAAAGTG	3969
b	4073	TTAAGTGGCTTTATGAGTCTCTTCAAGAACATTTCTCTGACGAGACAGTGAAGAAAGTG	4132
Y	3970	GAATATGATCATGAAACTCAACCTTATTTAGATTTCTGTAAATGAAACCACTTTTGT	4029
b	4113	GAATATGATCATGAAACTCAACCTTATTTAGATTTCTGTAAATGAAACCACTTTTGT	4192
Y	4030	GTTAATATTTCTGCTGTGAGAACTTTAAAGTTCAATACCAGATGCTTCAAGTCTCC	4089
b	4193	GTTAATATTTCTGCTGTGAGAACTTTAAAGTTCAATACCAGATGCTTCAAGTCTCC	4252
Y	4090	ATATGATGATCATGAGCAAGAGACAGAGCGGTGAGAAATTAACCTTCAAGTGAAG	4149
b	4253	ATATGATGATCATGAGCAAGAGACAGAGCGGTGAGAAATTAACCTTCAAGTGAAG	4312
Y	4150	CTGTCTCTCTGTGACCTTTTCAAGTATGTCAGAGGCTGCTGCTTGTGAGATGAGCT	4209
b	4313	CTGTCTCTCTGTGACCTTTTCAAGTATGTCAGAGGCTGCTGCTTGTGAGATGAGCT	4372
Y	4210	TCAAGCTCCCATATATCACTTCAATGATTTTATTTCTGTTCAAGCTTCTGACTTT	4269
b	4373	TCAAGCTCCCATATATCACTTCAATGATTTTATTTCTGTTCAAGCTTCTGACTTT	4432
Y	4270	ATGGAACCTTGGCTGTGA	4287
b	4433	ATGGAACCTTGGCTGTGA	4450

## RESULT 2

AF410459 5883 bp mRNA linear PRI 02-MAR-2002

DEFINITION Homo sapiens CD109 (CD109) mRNA, complete cds.  
 ACCESSION AF410459  
 VERSION AF410459.1 GI:19071208  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 5883)  
 AUTHORS Lin, M., Sutherland, D.R., Horsfall, W., Toty, N., Yeo, E., Nayar, R.,  
 Wu, X.-F. and Schuh, A.C.  
 TITLE Direct Submission  
 JOURNAL Submitted (14-AUG-2001) Medicine, University of Toronto, 1 King's  
 College Circle, Room 7366, Toronto, Ontario M5S 1A8, Canada

CELL surface antigen CD109 is a novel member of the alpha(2)  
 macroglobulin/C3, C4, C5 family of thioester-containing proteins  
 Blood 99 (5), 1683-1691 (2002)

JOURNAL MEDLINE 11861284  
 PUBMED 27079742

REFERENCE 2 (bases 1 to 5883)  
 AUTHORS Lin, M., Sutherland, D.R., Horsfall, W., Toty, N., Yeo, E., Nayar, R.,  
 Wu, X.-F. and Schuh, A.C.  
 TITLE Direct Submission  
 JOURNAL Submitted (14-AUG-2001) Medicine, University of Toronto, 1 King's  
 College Circle, Room 7366, Toronto, Ontario M5S 1A8, Canada

FEATURES  
 source 1..5883  
 /organism="Homo sapiens"

1..5883  
 /organism="Homo sapiens"

gene	5'UTR	CDS		/mol_type="mRNA"	/db_xref="taxon:9606"	/clone="K1"	1..5883	/gene="CD109"	1..112	/gene="CD109"	113..4450	/note="CD109"	/note="associated with the Gov alloantigen system"	/codon_start=1	/product="CD109"	/protein_id="AA84159.1"	/db_xref="GI:19071209"	/translation="MOGPELTLAHLVCVTAALAAVAPGPRVLVTPAGIIRPGANTV GVLEHCPSOVYKAEILKTAENLTVSLVLEAGVEKESFKTLPLSPLENSADELY ELRVTRGRODELIFNSRLSPETKLSVFIQDKALYKPKOKRKRIVLTFSDRFRY KSNAILIDPXSNILOQMLSOOSLDGSKTFOLSHSLILGMSLOVQNDQTYOS FOVSEYVLPFEVTLTQPLCYKSNRSHAGCTITAKTYKPKYKGVDTLPLSPFGK KNVITTFKINSANSPFDEKKNVDSNGLSEYLDSPPEVILITVTRESYVGI SRVSTNVFVKOHDIIEFDDYTVLPSLNTATVATVTRADGNOLITERRNNVIT VTQNTYEVWGSNGSNQKMEAVOKINVTYVPGSTFKIEPILIDESSLOLKVPLGS KSNMAYHSLEKSPSKTYIOLKTRDENIKVGSPELVSGNRLKELSVYVSGOLVA VGNQNTMBRLTPENSWTBAQVYVYIDEDGELISDVLKIPOLYKKNKIKLYMSKV KAPSEKVSRLISVTPDSIVGIVADKSNINMANSDITMENVAHLELHYNGYVLG MFNNSFAVFOECGLMWLTDMNLTNDYIDVYDAAERPMENREHVIDHPSIGS SPHVRKHPETWIMLDMNGRYIOBEVTPDSITSWATGVISBDSLGALTTTPV ELGAPQPFIFLNLPSVIRGESEFALIEITFNLYKDAIEVKVIEKSDKDLMTSE INANGHOOTLPSSEDAVLPPIRPHLGEIPIVYALSPASDAVQWILVAKES EKYSQSLIDLDIDNHLQSTLKTLSFSPNPVTGSEKVOITAGVLGSGINGASL IRMEYVGEQNMNFANITLIDYLRKKOITJNKLKELSPMOGORELLOREDG SFSAPGNDVDSGFWLSAFVLRPFLEADPIDIDONVLRHTTTLKHOQSNENPDP LAVITVLSQGNKSPVLTAVIYVSLGKRYOPNDVQSHIFLESERGLDI EVA LAVITVLSQGNKSPVLTAVIYVSLGKRYOPNDVQSHIFLESERGLDI EVA YALSHLOPOTSEKPIRMEHLSRORSLSGFASTOTVYALALSEFALMETERN IQVTVTPSSPSPKPKLIDPHNLLIOTLAVQVPAVYVANSNGFGLCOLNVVN VKAGSSRRRSRSTIONQAPLIDVAVENKODIANHNVCTSGSGPGRGALAEVNL LSGFMVSEPAISLSETVKVEYDGLNLTLDVSNENFQCNIPAYNPKYSNODLS VSIQVYVEPRQAVRSYNSVEKLSDDLSDVQGRCEDEGASGHHSSVIFPCR LLYFMELMWL"																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	
BASE COUNT	1697 a	1173 c	1241 g	1772 t																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																															</

## BASE COUNT

1697 a 1173 c 1241 g 1772 t

## ORIGIN

Query Match 98.5%; Score 4224.4; DB 9; Length 5883;  
 Best Local Similarity 98.8%; Pred. No. 0;  
 Matches 4286; Conservative 0; Mismatches 1; Indels 51; Gaps 1;

Y	1	ATGCGAGGCGCCACGCTCTGACCGCGCCGACCTCTCTGCTGTGACCGCGCGCTG	60
b	113	ATGCGAGGCGCCACGCTCTGACCGCGCCGACCTCTCTGCTGTGACCGCGCGCTG	172
Y	61	GCGTGGCTCCCGGCGCTGCTTCTGTGTGACGAGCCGAGATCATGAGCCCGAGGA	120
b	173	GCGTGGCTCCCGGCGCTGCTTCTGTGTGACGAGCCGAGATCATGAGCCCGAGGA	232
Y	121	AATGTACTATTGGGGTGAAGCTTCTGGAACACTGCGCTTCAAGGTGCTGGAAGGGG	180
b	233	AATGTACTATTGGGGTGAAGCTTCTGGAACACTGCGCTTCAAGGTGCTGGAAGGGG	292
Y	181	GAGCTCTCAAGACACATCAACCTCACTGCTCTGTCTGTGAGACAGAGAGTCTTT	240
b	293	GAGCTCTCAAGACACATCAACCTCACTGCTCTGTCTGTGAGACAGAGAGTCTTT	352
Y	241	GAAGAAAGGCTCTTTAAGACACTTACTCTTCATCACTTCACTTGAACAGTGAAGTGG	300
b	353	GAAGAAAGGCTCTTTAAGACACTTACTCTTCATCACTTCACTTGAACAGTGAAGTGG	412
Y	301	ATTATGAGCTACGCTAACCGAGCTACCCAGATGAGATTTTATCTCTATATAGTACC	360



**THIS PAGE BLANK (USPTO)**

```
Db 413 AATTATGAGCTAGCTGTAACCGGACGTAACCGAGAGAGATTTATCTTAACTAGTACC 472
Qy 361 CGCTTATCATTTTGAAGACCAAGAGATATCTGCTTCAATTCACACAGAGGCTTAAGC 420
Db 473 CGCTTATCATTTTGAAGACCAAGAGATATCTGCTTCAATTCACACAGAGGCTTAAGC 532
Qy 421 AAGCCAAAGCAAGAGAGTGAAGTTGCGATTTGATCACTCTTCCAGATTTTAAGCTTAC 480
Db 533 AAGCCAAAGCAAGAGAGTGAAGTTGCGATTTGATCACTCTTCCAGATTTTAAGCTTAC 592
Qy 481 AAAACCTCTTAAACATCTTCAATTAAGAGACCCCAATTCAAATTTGATCCACAGTGTG 540
Db 593 AAAACCTCTTAAACATCTTCAATTAAGAGACCCCAATTCAAATTTGATCCACAGTGTG 652
Qy 541 TCACAAACAAGTATCTTGAAGTCACTTCCAAAACCTTTCAGCTATCTCCATCCAAAT 600
Db 653 TCACAAACAAGTATCTTGAAGTCACTTCCAAAACCTTTCAGCTATCTCCATCCAAAT 712
Qy 601 CTGCTGATCTGCTTATTCAGTTCAGTGAAGTGAACAGACATATTAATCAATTCATTCAG 660
Db 713 CTGCTGATCTGCTTATTCAGTTCAGTGAAGTGAACAGACATATTAATCAATTCATTCAG 772
Qy 661 GTTTCGAATATGATTTACCAAAATTTGAAGTCACTTTCAGACACCATTAATTTGTTCT 720
Db 773 GTTTCGAATATGATTTACCAAAATTTGAAGTCACTTTCAGACACCATTAATTTGTTCT 832
Qy 721 ATGAATTTAGATTTAAATGATACCAATGAGGCAAGATATCAATATGAGGCAAGT 780
Db 833 ATGAATTTAGATTTAAATGATACCAATGAGGCAAGATATCAATATGAGGCAAGT 892
Qy 781 AAGAGAGAGTACGCTTACCTTTTACCTTATCTTTTGGGGAAGAGAAAATTT 840
Db 893 AAGAGAGAGTACGCTTACCTTTTACCTTATCTTTTGGGGAAGAGAAAATTT 952
Qy 841 ACAAACAATTTAAATATGATGATGCAAACTTCTCTTTTAAATGATGAAGATGAA 900
Db 953 ACAAACAATTTAAATATGATGATGCAAACTTCTCTTTTAAATGATGAAGATGAA 1012
Qy 901 AATGATGATGATCTTCAAAATGAGCTTCTGAAATCTGGAATCTATCTTCCCTGAGCA 960
Db 1013 AATGATGATGATCTTCAAAATGAGCTTCTGAAATCTGGAATCTATCTTCCCTGAGCA 1072
Qy 961 GTAGAAATTTTAAACAGAGTGAAGATCAATCAAGGATTTTCAAGAAATGAGAGCT 1020
Db 1073 GTAGAAATTTTAAACAGAGTGAAGATCAATCAAGGATTTTCAAGAAATGAGAGCT 1132
Qy 1021 AATGATGATGATCTTCAAGCAATGATTAATCAATGATTTTGAATTAATTAATCTTGG 1080
Db 1133 AATGATGATGATCTTCAAGCAATGATTAATCAATGATTTTGAATTAATTAATCTTGG 1192
Qy 1081 AAGCATCTCTCAATTCACAGGCACTGTAAGAGTAACTCTGCTGATGAGCAACCACTG 1140
Db 1193 AAGCATCTCTCAATTCACAGGCACTGTAAGAGTAACTCTGCTGATGAGCAACCACTG 1252
Qy 1141 ACTCTTGAAGAAAGAGAAATATATGATGATTAAGTGAAGACAGAGAACTAATAGAG 1200
Db 1253 ACTCTTGAAGAAAGAGAAATATATGATGATTAAGTGAAGACAGAGAACTAATAGAG 1312
Qy 1201 TACTGAGAGCGATCTTAACAGTGAAGATCAAGAAATGAGAGCTGTTCAAGAAATTAAT 1260
Db 1313 TACTGAGAGCGATCTTAACAGTGAAGATCAAGAAATGAGAGCTGTTCAAGAAATTAAT 1372
Qy 1261 ACTGCCCCCAAGGAGAACTTTAAAGTGAATTCCTCAATCTGAGAGATTCAGAGAG 1320
Db 1373 ACTGCCCCCAAGGAGAACTTTAAAGTGAATTCCTCAATCTGAGAGATTCAGAGAG 1432
Qy 1321 CTACAGTTGAAGGCTTATTTCTTGTGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1380
Db 1433 CTACAGTTGAAGGCTTATTTCTTGTGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1492
Qy 1381 AAGTCTCTAGTAAGACATCACTCACTAAACCAAGATGAAGAAATATTAAGGTGAG 1440
Db 1493 AAGTCTCTAGTAAGACATCACTCACTAAACCAAGATGAAGAAATATTAAGGTGAG 1552
Qy 1441 TCGCTTTTGAAGTGTGATGAGTGAAGCAACCAAGATTAAGTGAAGTGAAGTGAAG 1500
Db 1553 TCGCTTTTGAAGTGTGATGAGTGAAGCAACCAAGATTAAGTGAAGTGAAGTGAAG 1612
Qy 1501 GTATCAGAGGAGCAGTGTGAGTGAAGAAACCAAAATTCACAAATGTTCTCTTAAACA 1560
Db 1613 GTATCAGAGGAGCAGTGTGAGTGAAGAAACCAAAATTCACAAATGTTCTCTTAAACA 1672
Qy 1561 CCAGAAATTTTGGACTCCAAAAGCCTGTAAATGATTAATTAATGAAGTGAAGTGAAG 1620
Db 1673 CCAGAAATTTTGGACTCCAAAAGCCTGTAAATGATTAATTAATGAAGTGAAGTGAAG 1732
Qy 1621 GAATTAATGATGATTTTAAATTTCTGTTGATGATGATGATGATGATGATGATGAT 1680
Db 1733 GAATTAATGATGATTTTAAATTTCTGTTGATGATGATGATGATGATGATGATGAT 1792
Qy 1681 CTATATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1740
Db 1793 CTATATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1852
Qy 1741 ACAGAGCTGACTCATAGTGTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1800
Db 1853 ACAGAGCTGACTCATAGTGTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1912
Qy 1801 GCTCTAATGATTAATTAACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1860
Db 1913 GCTCTAATGATTAATTAACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1972
Qy 1861 TATTAATTAAGCAGTTCAGATTTCTTTCAGTCTTTCAGATGATGATGATGATGATGAT 1920
Db 1973 TATTAATTAAGCAGTTCAGATTTCTTTCAGTCTTTCAGATGATGATGATGATGATGAT 2032
Qy 1921 TTGACAGATGCAAACTCAGAGAGATTAATTAATGATGATGATGATGATGATGATGAT 1980
Db 2033 TTGACAGATGCAAACTCAGAGAGATTAATTAATGATGATGATGATGATGATGATGAT 2092
Qy 1981 GCTGAGAGTGTATGAGAGAAATGAAGACATATTTGATGATGATGATGATGATGATGAT 2040
Db 2093 GCTGAGAGTGTATGAGAGAAATGAAGACATATTTGATGATGATGATGATGATGATGAT 2152
Qy 2041 GGTGAGAGTCCACATGCTCCAGAAACATTTTCCAGAGCTTGAATTTGCTAGACCAAC 2100
Db 2153 GGTGAGAGTCCACATGCTCCAGAAACATTTTCCAGAGCTTGAATTTGCTAGACCAAC 2212
Qy 2101 ATGGGTTACAGAGATTAACCAAGATTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2160
Db 2213 ATGGGTTACAGAGATTAACCAAGATTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2272
Qy 2161 GTGGCTAATGATGATCTGAGAGACCTGGGCTTGGATTAACCAATCTCCAGT 2220
Db 2273 GTGGCTAATGATGATCTGAGAGACCTGGGCTTGGATTAACCAATCTCCAGT 2332
Qy 2221 GAGTCCAGAGCTTCCACATTTTCAATTTTGAATTTTGAATTTTGAATTTTGAATTTGA 2280
Db 2333 GAGTCCAGAGCTTCCACATTTTCAATTTTGAATTTTGAATTTTGAATTTTGAATTTGA 2392
Qy 2281 GGTGAAGAAATTTCTTGAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2340
Db 2393 GGTGAAGAAATTTCTTGAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2452
Qy 2341 AAGGTAATCATTAAGAAAGTGAAGAAATTTGAATTTGAATTTGAATTTGAATTTGAATTTGA 2400
Db 2453 AAGGTAATCATTAAGAAAGTGAAGAAATTTGAATTTGAATTTGAATTTGAATTTGAATTTGA 2512
Qy 2401 GCACAGGCAACAGAGAGACCTTCTGCTTCCAGTGAAGATGAGGCAACCTGTTCTTTT 2460
Db 2513 GCACAGGCAACAGAGAGACCTTCTGCTTCCAGTGAAGATGAGGCAACCTGTTCTTTT 2572
Qy 2461 CCATCAGGCAACATCTGAGAGAAATTTCTATCAGAGTGAAGTGAAGTGAAGTGAAGTGAAG 2520
Db 2573 CCATCAGGCAACATCTGAGAGAAATTTCTATCAGAGTGAAGTGAAGTGAAGTGAAGTGAAG 2632
```

**THIS PAGE BLANK (USPTO)**



**THIS PAGE BLANK (USPTO)**

RESULT 6  
AX537610  
LOCUS AX537610 4335 bp DNA linear PAT 23-NOV-2002  
DEFINITION Sequence 1 from Patent WO02070738.  
ACCESSION AX537610  
VERSION AX537610.1 GI:25269531  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE 1  
AUTHORS Schuh, A. and Ouwehand, W.  
TITLE Diagnosis and treatment of blood disorders  
JOURNAL Patent: WO 02070738-A 1 12-SEP-2002;  
Schuh, *Ante* *Ante*  
FEATURES  
source  
CDS  
1..4335  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
1..4335  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD57496.1"  
/db\_xref="GI:25269532"  
/translation="MOGPELITANLHLCVTAALAVAPGRPLVTPAGIIRPGANTYI  
GTELEHPCQVTVKALKTANLTVSVLEAGVTEKSPKLTPLPSLPLNADRIY  
ELAVTGRTODEILFNSNTRLSFETKRSVFIQDKALYRKQKRIYVLPEDPDY  
KTSNLNLIKDPKSNLIQOOSDLGKISFTQSLPILGWSIJOVNDTYS  
FQVSEVLPKPEVTLQPLVCSNMSKMLGITAKYVGRVGDVTLPLPSFWGK  
KKNITKFKINGSANFSPNDBEMKUNWDSNGISEYDLSPPGVEILLTVTESVGI  
SRVSTNVFPRKQNDYIIEFDYTVLPKPLNTPRTVTPADNQLTEERRNNVYT  
VTQNTAEYWSGNSGQKMAVQKINIVPQSTFKIEBPILIESEELQKAYPLGS  
KSSMAVSLFKPSKTYIOLKTRDENIKVSPFELVSGNRLEKELSYMYVSGOLVA  
VGKQNTMELSTPNSWTPKACIVIVYIIEDEGEIISVLKIPVOLVKNKIKLMSYV  
KAERSKVSLEISTQPDSTVIGIVAVDKSVLANASNDITMENVHLELYNTGYLG  
MENNFAVQECGMVLTDAITQYIDGVNDAEYARFEBENEGHIVIDHLSIGS  
SPVRAKHEPPELWLDNMGRIYQEPVTPDSTSVAVAGFVIEKSDKFLIMTNS  
ELQAFQPFPELNLPSVIRGEBEPALLETIRNLIKDAVEVYIIEKSDKFLIMTNS  
INATGHOQLTLVPSDEGATVLEPIRPHLGEIPIYTVTLSTPADAATOMLVAAEGL  
EKYSOSILDLDRLOSTLTKISFSPPNVTSSEVOITAGIDVPSINLAEL  
IRMPYCGEOMINPAPNIVYIDVLYLKKOULDNKEALSPMGQYORELQYQREDS  
SFAFGNDPDSGWTLSAFLRCFLEADPIDIDONVLRITYMLKQKNGEWD  
GRVHSELQGNKSPVTVTAYIVSLIGRYKQPNIDVQESIHFESEFSGISDNVT  
LALITVALSVSPRAKEALMLTPRABEGMOPWSESLSMSQPSRLDIEVAA  
VALLSHFLOFQSEBGPIMRWLSRQNSLGSASTODTTVALKALSEFALMTERTN  
VQVTVGSSPSPVKFLIDTNRLLQTELELAVOVPMANVCTSPSGRGAALMENVL  
VQVSSRRRSIIONOEAFDLVAVENKEDLNHVDLVNCTSPSGRGAALMENVL  
LSGFWPSRAISLSTVKKVEYDHGKLNLYLDSVNETQCVNIPAVRNKSVNTQDAS  
VSLVDYEPKQKRSYNSBVLKSDLSVDVQGRPCDASGSHHSVIFICFK  
LALYFELML"

COUNT 1296 a 837 c 922 g 1280 t

98.5%; Score 4221.4; DB 6; Length 4335;  
Similarity 98.8%; Pred. No. 0;  
Conservative 0; Mismatches 1; Indels 51; Gaps 1;

1 ATGCAAGGCGCCACCGGCTCTGACCGCGCCACCTCTCTGCTGTCACCGCGCGCTG 60  
1 ATGCAAGGCGCCACCGGCTCTGACCGCGCCACCTCTCTGCTGTCACCGCGCGCTG 60  
61 GCCCGGCTCCCGGCGCTCGGTTCTGCTGTCACCGCGCCACCGGCGCGCTG 120  
61 GCCCGGCTCCCGGCGCTCGGTTCTGCTGTCACCGCGCCACCGGCGCGCTG 120  
121 AATGAGCTATTTGGGGTGAAGCTTTGGAAACATCGCCCTTCAAGGTGACTGAAAGCG 180  
121 AATGAGCTATTTGGGGTGAAGCTTTGGAAACATCGCCCTTCAAGGTGACTGAAAGCG 180  
181 GAGCTGCTCAAGACAGATCAACCTCACTGCTCTGCTGCTGAAAGAGAGAGCTTT 240  
181 GAGCTGCTCAAGACAGATCAACCTCACTGCTCTGCTGCTGAAAGAGAGAGCTTT 240  
241 GAAAAAGCTCTTTAAGACACTTACTCTTCATCATCTCTGAAACAGTGCAGATGAG 300  
241 GAAAAAGCTCTTTAAGACACTTACTCTTCATCATCTCTGAAACAGTGCAGATGAG 300  
301 ATTATGAGCTACGTGTAAACCGAGCTACCCAGGATGAGATTTATTTCTTAATAGTACC 360  
301 ATTATGAGCTACGTGTAAACCGAGCTACCCAGGATGAGATTTATTTCTTAATAGTACC 360  
361 CGCTTATCATTTGAGACCAAGAAATATCTGCTCATTTCAACAGACAGGCTTTATAC 420  
361 CGCTTATCATTTGAGACCAAGAAATATCTGCTCATTTCAACAGACAGGCTTTATAC 420  
421 AAGCCAAAGCAAGAGTGAAGTTTCGATTGTACACTCTTCTCAAGTTTAAAGCTTAC 480  
421 AAGCCAAAGCAAGAGTGAAGTTTCGATTGTACACTCTTCTCAAGTTTAAAGCTTAC 480  
481 AAAACCTCTTAAACATTTCTATTAAGAACCCCAATCAATTTGATCCAAAGTGTG 540  
481 AAAACCTCTTAAACATTTCTATTAAGAACCCCAATCAATTTGATCCAAAGTGTG 540  
541 AAAACCTCTTAAACATTTCTATTAAGAACCCCAATCAATTTGATCCAAAGTGTG 540  
541 AAAACCTCTTAAACATTTCTATTAAGAACCCCAATCAATTTGATCCAAAGTGTG 540  
541 TCACAAACAAAGATCTTGAGTCAATTTCCAAACCTTTTCAAGTATCTCCCATCAATA 600  
541 TCACAAACAAAGATCTTGAGTCAATTTCCAAACCTTTTCAAGTATCTCCCATCAATA 600  
601 CTTGGTGAAGTGTCTATTCAGATTCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 660  
601 CTTGGTGAAGTGTCTATTCAGATTCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 660  
661 GTTTCAGAAATATGATTAACCAAAATTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 720  
661 GTTTCAGAAATATGATTAACCAAAATTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 720  
721 ATGAATTCATGAATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTA 780  
721 ATGAATTCATGAATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTA 780  
781 AAGGAGAGCTAAGCTTACATTTTACCTTTTGGGAAAGAAAGAAATATTT 840  
781 AAGGAGAGCTAAGCTTACATTTTACCTTTTGGGAAAGAAAGAAATATTT 840  
841 AAAAAACATTTAAGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTA 900  
841 AAAAAACATTTAAGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTA 900  
901 AATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATG 960  
901 AATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATG 960  
961 GTAGAAATTTTAAACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1020  
961 GTAGAAATTTTAAACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1020  
1021 AATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATG 1080  
1021 AATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATG 1080  
1081 AAGCATCTCTCAACTTCAAGCAAGCTGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1140  
1081 AAGCATCTCTCAACTTCAAGCAAGCTGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1140

**THIS PAGE BLANK (USPTO)**

D	2161	GTGGCTACTGGTTTGTGTGATCTCTGAGAGACTGGGCTTGGACATTAACAATACTACCGAGTG	2220
Q	2221	GAGCTCCAGACCTTCCAAACCAATTTTTCATTTTGTGAATCTTCCCTACTGTGTATCAGA	2280
D	2221	GAGCTCCAGACCTTCCAAACCAATTTTTCATTTTGTGAATCTTCCCTACTGTGTATCAGA	2280
Q	2281	GGTGAAGAAATTTGCTTTGGAAAATACATATATTCATTTATTTGGAAGATGCCACTGAGGT	2340
D	2281	GGTGAAGAAATTTGCTTTGGAAAATACATATATTCATTTATTTGGAAGATGCCACTGAGGT	2340
Q	2341	AAGGTAATCATGAGAAAAAGTGCAAAATTTGATTTCTAATGACCTTCAAGTGAATTAAT	2400
D	2341	AAGGTAATCATGAGAAAAAGTGCAAAATTTGATTTCTAATGACCTTCAATGAAATTAAT	2400
Q	2401	GCCACAGGCCAACCCAGACGCCCTTCGGTCTCCAGTGAAGATGGGGCAACTGTTCTTTT	2460
D	2401	GCCACAGGCCAACCCAGACGCCCTTCGGTCTCCAGTGAAGATGGGGCAACTGTTCTTTT	2460
Q	2461	CCCATCAGGCCCAACATCTGGAGAAAATCTCTATCAGTCACAGCTCTTTCACCCACT	2520
D	2461	CCCATCAGGCCCAACATCTGGAGAAAATCTCTATCAGTCACAGCTCTTTCACCCACT	2520
Q	2521	GCTTCTGATGCTGTACCCCAATGATTTTAAAGGCTGAAGAAATGAAAAATCATAT	2580
D	2521	GCTTCTGATGCTGTACCCCAATGATTTTAAAGGCTGAAGAAATGAAAAATCATAT	2580
Q	2581	TCACAATTCATCTTATTTAGACTGACGAAATAGGCTTACAGAGTACCTGAAAACCTTG	2640
D	2581	TCACAATTCATCTTATTTAGACTGACGAAATAGGCTTACAGAGTACCTGAAAACCTTG	2640
Q	2641	AGTTTCTCATTTCTCTCTAATATACAGTGCATGGCAGTGAAGATTCAGATCACGCAATT	2700
D	2641	AGTTTCTCATTTCTCTCTAATATACAGTGCATGGCAGTGAAGATTCAGATCACGCAATT	2700
Q	2701	GGAGATGTTCTTGGTCTCTCCATCAATAGGCTTACCTCATTTGATTCGATTCCTTATGGC	2760
D	2701	GGAGATGTTCTTGGTCTCTCCATCAATAGGCTTACCTCATTTGATTCGATTCCTTATGGC	2760
Q	2761	TGTGTGAACAGAACATGATTAATTTTGTCTCCAAATATTTACATTTTGGATTTATCTGACT	2820
D	2761	TGTGTGAACAGAACATGATTAATTTTGTCTCCAAATATTTACATTTTGGATTTATCTGACT	2820
Q	2821	AAAAAGAAACAATGACAGATTAATTTTGAAGAAAAAAGCTCTTTCATTTATAGGCAAGT	2880
D	2821	AAAAAGAAACAATGACAGATTAATTTTGAAGAAAAAAGCTCTTTCATTTATAGGCAAGT	2880
Q	2881	TACCAAGAGAACTTCTATCAGAGGGAAGATGGCTTTCAAGTCTTTTGGGAATTAAT	2940
D	2881	TACCAAGAGAACTTCTATCAGAGGGAAGATGGCTTTCAAGTCTTTTGGGAATTAAT	2940
Q	2941	GACCCCTTCTGGAGCACTTGGTTGTCAAGCTTTGTTTAAAGATGTTTCTTGAAGCCGAT	3000
D	2941	GACCCCTTCTGGAGCACTTGGTTGTCAAGCTTTGTTTAAAGATGTTTCTTGAAGCCGAT	3000
Q	3001	CCTTACATGATATTTGATCAGAAATGTGTACACAGAACTACACTTGGCTTAAAGGACAT	3060
D	3001	CCTTACATGATATTTGATCAGAAATGTGTACACAGAACTACACTTGGCTTAAAGGACAT	3060
Q	3061	CAGAAATCCACGGTGAATTTTGGGATCCAGGAAGATGATTCATATGTAAGCTTCAAGT	3120
D	3061	CAGAAATCCACGGTGAATTTTGGGATCCAGGAAGATGATTCATATGTAAGCTTCAAGT	3120
Q	3121	GGCAATTAAGATCCAGTACATTAACAGCTATATTTGTAATCTTCTCCTGGGATATAGA	3180
D	3121	GGCAATTAAGATCCAGTACATTAACAGCTATATTTGTAATCTTCTCCTGGGATATAGA	3180
Q	3181	AAGTATCAGCCTTAATGATGTGCAAGTCAATCCATTTTGGAGTCTGAATTCAGT	3240
D	3181	AAGTATCAGCCTTAATGATGTGCAAGTCAATCCATTTTGGAGTCTGAATTCAGT	3240
Q	3241	AGAGGAATTTACACAATTTATCTAGCCCTTAACTTATGCACTGACATCAATGGGG	3300
D	3241	AGAGGAATTTACACAATTTATCTAGCCCTTAACTTATGCACTGACATCAATGGGG	3300



**THIS PAGE BLANK (USPTO)**

Qy	3301	AGTCCTAAAGGAAAGAAAGCTTTAAATATGCTGACTTGGAGAGAGAAACAAGAAAGTGGC	3360
Dp	3301	AGCTCTTAAGCGAAGAAAGCTTTGAATATGCTGACTTGGAGAGAGAAACAAGAAAGTGGC	3360
Qy	3361	ATGCATATCTGGGTGTCATCAGAGTCCAACTTTCTGACTCTGGACGCCAGCTCCCTG	3420
Dp	3361	ATGCATATCTGGGTGTCATCAGAGTCCAACTTTCTGACTCTGGACGCCAGCTCCCTG	3420
Qy	3421	GATATTGAAGTTGCAGCCTATGCACTGCTCTCAACTTCTTAACAATTCAGACTTTGAG	3480
Dp	3421	GATATTGAAGTTGCAGCCTATGCACTGCTCTCAACTTCTTAACAATTCAGACTTTGAG	3480
Qy	3481	GGAATCCCAATTAGAAGGTGGCTAGACAGGCAAGAAATAGCTGGGGGTGTTTGCACT	3540
Dp	3481	GGAATCCCAATTAGAAGGTGGCTAGACAGGCAAGAAATAGCTGGGGGTGTTTGCACT	3540
Qy	3541	ACTCAGATACCACTGTGGCTTTAAAGGCTGTCTGAAATTTGAGAGCCCTATAGATAA	3600
Dp	3541	ACTCAGATACCACTGTGGCTTTAAAGGCTGTCTGAAATTTGAGAGCCCTATAGATAA	3600
Qy	3601	GAAAGACAAATATTCAGATGACCGTGAACGGGCTTACTCAACAATCTT-----	3651
Dp	3601	GAAAGACAAATATTCAGATGACCGTGAACGGGCTTACTCAACAATCTT-----	3651
Qy	3652	-----CTTGGCTGTGGTACAGCCA	3659
Dp	3651	CTGATTGACACACACACACCCTTACTCTTCAGACAGCAGAGCTTGTGTGTACAGCCA	3720
Qy	3670	ATGCGATTAATATTTCCGCAAAATGGTTTGGATTTGCTATTTGTCACTCAATGTGTGA	3729
Dp	3721	ATGCGATTAATATTTCCGCAAAATGGTTTGGATTTGCTATTTGTCACTCAATGTGTGA	3780
Qy	3730	TATATATGGAAGGCTTCGGGCTCTTGAAGAAGCAATCTATCCAAATCAAGAAAGCC	3789
Dp	3781	TATATATGGAAGGCTTCGGGCTCTTGAAGAAGCAATCTATCCAAATCAAGAAAGCC	3840
Qy	3790	TTTGAATTAAGATGTTGCTGTAAAGAAATAAAGATGATCTCAATCATGTGGATTTGAT	3849
Dp	3841	TTTGAATTAAGATGTTGCTGTAAAGAAATAAAGATGATCTCAATCATGTGGATTTGAT	3900
Qy	3850	GTTGTGTACAAGCTTTTCGGGCCCGGGTGAAGGTGCATGGCTCTTATGAAAGTTAAGCTA	3909
Dp	3901	GTTGTGTACAAGCTTTTCGGGCCCGGGTGAAGGTGCATGGCTCTTATGAAAGTTAAGCTA	3960
Qy	3910	TTAAGGGCTTATATGTTGCTTCAAGACAAATTTCTGAGGGGAGACAGTGAAGAAAGTG	3969
Dp	3961	TTAAGGGCTTATATGTTGCTTCAAGACAAATTTCTGAGGGGAGACAGTGAAGAAAGTG	4020
Qy	3970	GAAATATGATCATGAAAACTCAACTCTATTTAGATTCGTAAATGAAGAACCAAGTTTGT	4029
Dp	4021	GAAATATGATCATGAAAACTCAACTCTATTTAGATTCGTAAATGAAGAACCAAGTTTGT	4080
Qy	4030	GTTAATATATCTGCTGTGTGAAGAACTTTAAAGTTTCAAAATACCAAGATGCTTCACTGCC	4089
Dp	4081	GTTAATATATCTGCTGTGTGAAGAACTTTAAAGTTTCAAAATACCAAGATGCTTCACTGCC	4140
Qy	4090	ATATGAGATTTACATATGAGCAAGAGACAGAGGGGGTGAAGATTAACAATGAAAGTAGAG	4149
Dp	4141	ATATGAGATTTACATATGAGCAAGAGACAGAGGGGGTGAAGATTAACAATGAAAGTAGAG	4200
Qy	4150	CTGTCTCTCTGTGACCTTTGCAAGTATGTCAGAGGCTGCCCTCTTGTGAAGATGAGACT	4209
Dp	4201	CTGTCTCTCTGTGACCTTTGCAAGTATGTCAGAGGCTGCCCTCTTGTGAAGATGAGACT	4260
Qy	4210	TCAGAGCTCCATCATCACTTTCOAGTCATTTTATTTTCTGTTTCAAGCTTCTGTACTTT	4269
Dp	4261	TCAGAGCTCCATCATCACTTTCOAGTCATTTTATTTTCTGTTTCAAGCTTCTGTACTTT	4320
Qy	4270	ATGGAACCTTTGGCTG 4284	
Dp	4335	ATGGAACCTTTGGCTG 4335	

AX537612  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
JOURNAL  
FEATURES  
source  
CDS

AX537612  
Sequence 3 from Patent WO02070738.  
AX537612  
AX537612.1 GI:25269535  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

1  
Schuh, A. and Ouweland, W.  
Diagnosis and treatment of blood disorders  
Patient: WO 02070738-A 3 12-SEP-2002;  
Schuh, Andre (CA).  
Location/Qualifiers  
1..4335  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
1..>4335  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD57497.1"  
/db\_xref="GI:25269536"  
/translation="MQGPRLITAAHLCTALAAVAPGRPLVTAAGIIRPGAVTII  
GVEILHECPQVTVKAEILKTASNLTVSVLEAGVPEKSPFKTLLPSLPLNSADELY  
ELRVTGRTOEILFNSNTRLSFETKRISVLIQTDKALYKQKEVKEFIVLPSDFKY  
KTSNLTIKDPSKNTLIQOOLMSQSDSLIKETKLSHPILIGMSLOVQVNDQYYGS  
FOVEYVLPRKEVTLTOPLYCAMSKLINSTIIAKTYGKPKADVLTFLPSFMGK  
KKNITFKFKINGSANFENDEEMKNVMSNGSLSEYDLSPPGEVILTVTESVGI  
SRNSTVPEFKQHDIIIEPDYTVTLKPSLNTRATVATVTAADNQLLEBRNNVIT  
VTQANTYETWMSGNSQKQKMAVQKINYTPQSGTFKIEPILDESELQKAFPLSS  
KSSNAVSLFSPSKITQLKTRDENIKVSPFELVSGNRLKELSVVVSQQLVA  
VGKNSWTFSLTPENSNTTPKACVIVYIIEDEGLISVLIPOVLKRNKIKLYMSVK  
KAESERVSLSLITQOPDSIVGIVAVDVSNTLVANASNDITMENVHELELVNTYIGL  
MFNDSFVFOECGLMTVDANLTQDYIDGVYDNEVLERFMEENEGHVIDHDSILSG  
SPHRRKPEPTFWILDTNMSGRIYQEFEEVTPDSITSMVATGVFISDELGLITTPP  
ELQAPQPEPTFWILPYVINGEPEPALITTFPNYIKDQTEVYVILKSDKPIILTSN  
INAGHQDTLTPESDEGATVLPFRPHLEGIPIVTALSTFASDAYOTMLVKAEGT  
EKSSQSLILDLTDNRLQSTLKITLSFSPNPVTVSERSVQITAGVILSPINGLASI  
IRMYGSGEQQMNFAPNIYILDTLTKKKQLTDMKEKALSFMQGRQRELLYREDDQ  
SFSAFNVDPGSGTWSLAFVRCLEADPYIDIDONVLRHYTTLKGKQNGEMWD  
GRVHSELQNGKSPVLTLYATVLSLIGYRCQPNIDVQESIHFLSEFSNGISDNYA  
LALITVALSSVSGSPAKAKBALMLTRABQEBGMFVMSSEKSDSNRPRLDIEVAL  
YALSHLFOQTSGEPIPTMRWLSRQKNSLGFSASTQDTVALKALSERPALNTERIT  
IOLVNTGFPSPSPVKFLIDTNRLILQTAELAAVQPAVNLISANGFGAILQNLVNY  
VKAGSSRRRRISSIONOAFDLDDVAIKENQDLNVDLVCTSFSGPFGSGLAMLEVIN  
LSGAMVSEASISLSETVKAYVDHGKNTLYDSVNETOFQCNIPAVRPFKVSNTDDAA  
VSIDYVPEPRQAAVRSYNSVFKLSCLDSDVQGCRRPEDDASGHHSSVIFIFCP  
LLYMEMLL"

BASE COUNT 1295 a 838 c 922 g 1280 t  
ORIGIN

Query Match 98.4%; Score 4219.8; DB 6; Length 4335;  
Best Local Similarity 98.4%; Pred. No. 0;  
Matches 4282; Conservative 0; Mismatches 2; Indels 51; Gaps 1

1 ATGAGGAGCCACCGCTCTTACCGCGCGCCACCTCTCTGCGTGTGACCGCGCGCGCTG 60  
1 ATGACGAGGCCACCGCTCTTACCGCGCGCCACCTCTCTGCGTGTGACCGCGCGCGCTG 60  
GCCGTGGCTCCCGGGCTCCGGTTCTGGTGAACGCCGACGAGTATAGAGCCGGAGGA 120  
GCCGGTCCCGGGCTCCGGTTCTGGTGAACGCCGACGAGTATAGAGCCGGAGGA 120  
AATGATATTATGGGTGAGACTTCTGGAACAACCTGCCTTACACAGGTGACTGTGAAGCG 180  
AATGATATTATGGGTGAGACTTCTGGAACAACCTGCCTTACACAGGTGACTGTGAAGCG 180  
GAGCTGCTCAAGACGATCAACCTCTACTCTCTGCTCTGGAAGCAAGAGAGTCTTT 240

PLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2004, 12:36:15 ; Search time 1017 Seconds  
(without alignments)  
11379.047 Million cell updates/sec

Title: US-10-020-095-3

Perfect score: 4287  
Sequence: 1 atgcagggccacgcgcctc.....ttatggaacttgctgtga 4287

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: N\_Geneseq\_19Jun03.\*  
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*  
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	4281.2	99.9	4369	25 AAD49434 Human r150 gene #1
2	4224.4	98.5	4761	24 ABQ79964 Human CD109 K1 pro
3	4224.4	98.5	5882	25 AAD49440 Human blood cell s
4	4224.4	98.5	5883	24 ABQ79966 Human CD109 K1-H7
5	4224.4	98.5	5883	25 AAD49435 Human r150 DNA #1
6	4222.8	98.5	4761	24 ABQ79965 Human CD109 K1 var
7	4222.8	98.5	5895	24 ABQ79967 Human CD109 K1-H7
8	4221.4	98.5	4335	24 AAL49815 Human platelet all

9	4219.8	98.4	4335	24 AAL49816 Human platelet all
10	4219.6	98.4	4473	21 AAA62010 Hydrophobic domain
11	4216.6	98.4	4335	21 AAA60199 Hydrophobic domain
12	4140.2	96.6	4146	25 AAD49436 Human blood cell s
13	4082.4	95.2	4197	25 AAD49437 Human r150 DNA #2
14	3187.4	74.4	3535	24 ABQ79968 Human CD109 K15 pr
15	3185.8	74.3	3535	24 ABQ79969 Human CD109 K15 pr
16	1017	23.7	1448	24 ABQ79970 Human polynucleoti
17	975.4	23.7	1448	24 ABQ79971 Human polynucleoti
18	812.2	18.9	1459	22 ABA08825 Human secreted pro
19	730.8	17.0	2403	24 ABV77330 Alpha-1 protease
20	541.6	12.6	875	20 AAX20449 Human secretase and
21	525	12.2	821	25 ABX08843 Angiogenesis-assoc
22	497.6	11.6	744	20 AAZ15521 Human gene express
23	331.6	7.7	690	20 AAZ13539 Human gene express
24	300	7.0	300	20 AAZ13539 Human gene express
25	272	6.3	354	25 ABX47702 Bovine EST associa
26	175	4.1	744	20 AAZ16725 Human gene express
27	174	4.1	4527	22 AAC30951 C. elegans alpha-2
28	174	4.1	4527	22 AAC30951 C. elegans alpha-2
29	172.6	4.0	4560	22 AAC30952 Drosophila melanog
30	171.6	4.0	2285	23 ABL12153 Drosophila melanog
31	152.2	3.6	3612	23 ABL12153 Drosophila melanog
32	143.8	3.4	6324	23 ABL19704 Drosophila melanog
33	127.2	3.0	5611	23 ABL20610 Drosophila melanog
34	127.2	3.0	6979	23 ABL2152 Drosophila melanog
35	126.8	3.0	4615	23 AAS92132 DNA encoding novel
36	126.8	3.0	4629	24 AAS92132 DNA encoding novel
37	126.8	3.0	5815	21 AAC74484 Human ORF39 P
38	126.8	3.0	5815	21 AAC74484 Human ORF39 P
39	125.2	2.9	4487	24 ABR2035 DNA encoding novel
40	125.2	2.9	4488	24 ABR2035 DNA encoding novel
41	120.8	2.8	4660	24 ABQ3897 Human alpha-2-macr
42	120.6	2.8	2608	24 AAL49817 Human pregnancy zo
43	119.8	2.8	4501	25 ABX70464 Human platelet all
44	119.4	2.8	1140	22 ABA06596 DNA encoding human
45	119.4	2.8	1140	22 ABA06596 Human cDNA SEQ ID

ALIGNMENTS

RESULT 1	
AAD49434	
ID: AAD49434 standard; DNA; 4369 BP.	
XX	
AC	AAD49434;
XX	
DT	24-MAR-2003 (first entry)
XX	
DE	Human r150 gene #1.
XX	
KW	Human; GPI-anchored TGF-beta1 binding protein; r150 protein; cancer;
KW	glycosylphosphatidylinositol; transforming growth factor-beta1;
KW	therapy; gene; ds.
XX	
OS	Homo sapiens.
XX	
PH	Key
FT	CDS
FT	26..4312
FT	Location/Qualifiers
FT	/*tag= a
FT	/product= "Human r150 protein #1"
FT	/transl_except= (pos:2132..2134, aa:Xaa)
FT	/note= "Xaa corresponds to Ser, Tyr"
XX	
PN	W0200285942-A2
XX	
XX	
PD	31-OCT-2002.
XX	
PF	24-APR-2002; 2002MO-CA00560
XX	
PR	24-APR-2001; 2001US-285713P.
PR	14-FEB-2002; 2002US-356163P.

**THIS PAGE BLANK (USPTO)**

XX (UTMC-) UNIV MCGILL.  
 PA Phillip A, Tam B;  
 PI WPI: 2003-09100/08.  
 DR P-PSDB; AAE32012.  
 XX Novel transforming growth factor (TGF)-beta 1 binding reagent which  
 PT comprises r150 protein which acts as accessory receptor of TGF-beta,  
 PT useful for negatively modulating TGF-beta activity, and thus for  
 PT treating cancer  
 XX  
 PS Claim 7; Page 91-93; 127pp; English.  
 XX The invention relates to novel transforming growth factor (TGF)-beta1  
 CC binding reagent which comprises a glycosylphosphatidyl inositol (GPI)-  
 CC anchored TGF-beta1 binding protein referred to as r150 which acts as  
 CC an accessory receptor of TGF-beta. The invention is used for negatively  
 CC modulating TGF-beta activity, and thus for treating conditions  
 CC characterised by overproduction of TGF-beta, such as cancer. Antisense  
 CC molecules of the invention are used for increasing TGF-beta availability  
 CC and increase graft success. The present sequence is a gene encoding  
 CC human r150 protein.  
 XX  
 SQ Sequence 4369 BP; 1301 A; 842 C; 932 G; 1293 T; 1 other;  
 Query Match 99.9%; Score 4281.2; DB 25; Length 4369;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 4283; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Db 626 CTGTGAGCTGCTATTCAGTTCAAGTGAATGACACAGACATATCATTCATTTGAG 685  
 Qy 661 GTTTGAAATATGATTAACAAATTTGAAAGTACCTTTGACGACACCATTAATTTGCTT 720  
 Db 686 GTTTCAGAAATATATTAACAAATTTGAAAGTACCTTTGACGACACCATTAATTTGCTT 745  
 Qy 721 ATGAATTTGAAAGTAAATGATGACATCAACGCAAAAGTATCATATGGAAGCAGTG 780  
 Db 746 ATGAATTTGAAAGTAAATGATGACATCAACGCAAAAGTATCATATGGAAGCAGTG 805  
 Qy 781 AAGGAGAGCTAACGCTTACATTTTACCTTTATCCTTTTGGGAAAGAAATATTT 840  
 Db 806 AAGGAGAGCTAACGCTTACATTTTACCTTTATCCTTTTGGGAAAGAAATATTT 865  
 Qy 841 ACAAATTTTAAAGTAAATGATGACATCAACGCAAAAGTATCATATGGAAGCAGTG 900  
 Db 866 ACAAATTTTAAAGTAAATGATGACATCAACGCAAAAGTATCATATGGAAGCAGTG 925  
 Qy 901 AATGTAATGATTTCTTCAATGATGACATTTGCAATATCTGATCTATCTCCCTGACCA 960  
 Db 926 AATGTAATGATTTCTTCAATGATGACATTTGCAATATCTGATCTATCTCCCTGACCA 985  
 Qy 961 GTTGAATTTTAAACAGTGAACAGATCAATGATTAAGTATTTCAAGAAATGATGACT 1020  
 Db 986 GTTGAATTTTAAACAGTGAACAGATCAATGATTAAGTATTTCAAGAAATGATGACT 1045  
 Qy 1021 AATGTTCTTCAAGACATGATTAATCATATGATTTTGGATTAATGATGATGATGAT 1080  
 Db 1046 AATGTTCTTCAAGACATGATTAATCATATGATTTTGGATTAATGATGATGATGAT 1105  
 Qy 1081 AAGCCATCTCTCAACTTCAACGACATGTAAGTAACTGCTGATGATGACCAACTG 1140  
 Db 1106 AAGCCATCTCTCAACTTCAACGACATGTAAGTAACTGCTGATGATGACCAACTG 1165  
 Qy 1141 ACTCTTGAAGAAAGAAATATGATGATTAAGTAACTGCTGATGATGACCAACTG 1200  
 Db 1166 ACTCTTGAAGAAAGAAATATGATGATTAAGTAACTGCTGATGATGACCAACTG 1225  
 Qy 1201 TACTGAGGAGATCTTCAAGTGAATCAAGAAATGGAAGTCTTCAAGAAATTAATTT 1260  
 Db 1226 TACTGAGGAGATCTTCAAGTGAATCAAGAAATGGAAGTCTTCAAGAAATTAATTT 1285  
 Qy 1261 ACTGTCCTTCAAGTGAATCTTGAATGATTAAGTAACTGCTGATGATGACCAACTG 1320  
 Db 1286 ACTGTCCTTCAAGTGAATCTTGAATGATTAAGTAACTGCTGATGATGACCAACTG 1345  
 Qy 1321 CTACAGTTGAAGGCTTATTTCTTGTAGTAAAGTGAATGATGATGATGATGATGAT 1380  
 Db 1346 CTACAGTTGAAGGCTTATTTCTTGTAGTAAAGTGAATGATGATGATGATGATGAT 1405  
 Qy 1381 AAGTCTCTTGAAGTGAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 1440  
 Db 1406 AAGTCTCTTGAAGTGAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 1465  
 Qy 1441 TCGCTTTTGAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATG 1500  
 Db 1466 TCGCTTTTGAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATG 1525  
 Qy 1501 GTATCAAGGAGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATG 1560  
 Db 1526 GTATCAAGGAGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATG 1585  
 Qy 1561 CCAGAAATTTCTTGAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 1620  
 Db 1586 CCAGAAATTTCTTGAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 1645  
 Qy 1621 GAAATTTAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1680  
 Db 1646 GAAATTTAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1705  
 Qy 1681 CTATATTGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1740  
 Db 1706 CTATATTGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1765

**THIS PAGE BLANK (USPTO)**

QY 1741 ACAAGCCTGACATCCAGTAGTGGAGTTGAGCTGTGACAAAGAGTGAATCTGATGAT 1800  
 DB 1766 ACACAGCCTGACATCCAGTAGTGGAGTTGAGCTGTGACAAAGAGTGAATCTGATGAT 1825  
 QY 1801 GCGCTTAT 1860  
 DB 1826 GCGCTTAT 1885  
 QY 1861 TATTTATTTAGGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920  
 DB 1886 TATTTATTTAGGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1945  
 QY 1921 TTGACAGATGCAAACTCAGAAAGATTAATGATGATGATGATGATGATGATGATGAT 1980  
 DB 1946 TTGACAGATGCAAACTCAGAAAGATTAATGATGATGATGATGATGATGATGATGAT 2005  
 QY 1981 GCTGAGAGGTTTATGAGAGAAATGAGAGACATATGATGATGATGATGATGATGATG 2040  
 DB 2006 GCTGAGAGGTTTATGAGAGAAATGAGAGACATATGATGATGATGATGATGATGATG 2065  
 QY 2041 GGTAGAGCTGACATGTCGAGAAAGATTTCCAGAGACCTGATGATGATGATGATGATG 2100  
 DB 2066 GGTAGAGCTGACATGTCGAGAAAGATTTCCAGAGACCTGATGATGATGATGATGATG 2125  
 QY 2101 ATGGGTTTACAGATTTTACAAAGATTTGAGATGATGATGATGATGATGATGATGATG 2160  
 DB 2126 ATGGGTTTACAGATTTTACAAAGATTTGAGATGATGATGATGATGATGATGATGATG 2185  
 QY 2161 GTGGCTATGCTGTTTGTGATCTGAGAGACCTGAGCTGATGATGATGATGATGATGATG 2220  
 DB 2186 GTGGCTATGCTGTTTGTGATCTGAGAGACCTGAGCTGATGATGATGATGATGATGATG 2245  
 QY 2221 GAGCTCAAGCCTTCCAAACATTTTCAATTTTGAATCTGATGATGATGATGATGATGATG 2280  
 DB 2246 GAGCTCAAGCCTTCCAAACATTTTCAATTTTGAATCTGATGATGATGATGATGATGATG 2305  
 QY 2281 GGTGAAGAAATTTGCTTGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2340  
 DB 2306 GGTGAAGAAATTTGCTTGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2365  
 QY 2341 AAGGTATATGATGAGAAAGTGAACAATTTGATGATGATGATGATGATGATGATGATGAT 2400  
 DB 2366 AAGGTATATGATGAGAAAGTGAACAATTTGATGATGATGATGATGATGATGATGATGAT 2425  
 QY 2401 GGCACAGGCGCACAGAGAGACCTTGTGCTCCAGTGAAGATGAGAGATGAGAGATGAG 2460  
 DB 2426 GGCACAGGCGCACAGAGAGACCTTGTGCTCCAGTGAAGATGAGAGATGAGAGATGAG 2485  
 QY 2461 CCCATCAGGCGCAACATCTGAGAGAAATTCCTATCAAGTCAAGCTCTTTCAACCACAT 2520  
 DB 2486 CCCATCAGGCGCAACATCTGAGAGAAATTCCTATCAAGTCAAGCTCTTTCAACCACAT 2545  
 QY 2521 GCTTCTGATGCTGTCACCCAGATGATTTAGTAAAGCTGAAGAAATGAGAAATCATAT 2580  
 DB 2546 GCTTCTGATGCTGTCACCCAGATGATTTAGTAAAGCTGAAGAAATGAGAAATCATAT 2605  
 QY 2581 TCACATCCATCTTATTAAGCTGACGACAAATAGGTATAGAGTACAGTACCTGAAACCTT 2640  
 DB 2606 TCACATCCATCTTATTAAGCTGACGACAAATAGGTATAGAGTACAGTACCTGAAACCTT 2665  
 QY 2641 AGTTTCTCATTTCTCTCTTAATAGAGTACGAGTGAAGAGTTCAGATCACTGCAAT 2700  
 DB 2666 AGTTTCTCATTTCTCTCTTAATAGAGTACGAGTGAAGAGTTCAGATCACTGCAAT 2725  
 QY 2701 GAGAGATTTCTTGCTCTTCAATCAATGCTTAAGCTCAATGATGAGATGCTTATGCT 2760  
 DB 2726 GAGAGATTTCTTGCTCTTCAATCAATGCTTAAGCTCAATGATGAGATGCTTATGCT 2785  
 QY 2761 TGTGTGAACAGAAATGATTAATTTTGTCTCCAAATTTTGAATTTGATGATGATGATGAT 2820  
 DB 2786 TGTGTGAACAGAAATGATTAATTTTGTCTCCAAATTTTGAATTTGATGATGATGATGAT 2845

QY 2821 AAAAAGAAACACTGACAGATTAATTTGAAGAAAGAAAGCTCTTTCATTTATGAGCAAGGT 2880  
 DB 2846 AAAAAGAAACACTGACAGATTAATTTGAAGAAAGAAAGCTCTTTCATTTATGAGCAAGGT 2905  
 QY 2881 TACCAGAGAACTTCTCTATACAGAGAAAGAGGCTCTTTCAGTCTTTTGGAAATTAAT 2940  
 DB 2906 TACCAGAGAACTTCTCTATACAGAGAAAGAGGCTCTTTCAGTCTTTTGGAAATTAAT 2965  
 QY 2941 GACCTTCTGGAGACCTTGTGTCAGCTTTTGTGTTTGAAGATTTCCGTGAAGCCGAT 3000  
 DB 2966 GACCTTCTGGAGACCTTGTGTCAGCTTTTGTGTTTGAAGATTTCCGTGAAGCCGAT 3025  
 QY 3001 CCTTACATGATTTGATCAGAAATGCTTTTACACAGAACTACACTTGGCTTAAAGACAT 3060  
 DB 3026 CCTTACATGATTTGATCAGAAATGCTTTTACACAGAACTACACTTGGCTTAAAGACAT 3085  
 QY 3061 CAGAAATCCAAAGGTAATTTTGGGATCCAGAGAAAGGATTCATATGAGCTTCAAGGT 3120  
 DB 3086 CAGAAATCCAAAGGTAATTTTGGGATCCAGAGAAAGGATTCATATGAGCTTCAAGGT 3145  
 QY 3121 GGCATTAATAAGTCCAGTAAACCTTACAGCTTAATTTGTAATCTCTCTGGGATATAGA 3180  
 DB 3146 GGCATTAATAAGTCCAGTAAACCTTACAGCTTAATTTGTAATCTCTCTGGGATATAGA 3205  
 QY 3181 AAGTATCAGCTTACATGATGTCAGAAAGTCTATTCATTTTGGAGTCTGAATCAAT 3240  
 DB 3206 AAGTATCAGCTTACATGATGTCAGAAAGTCTATTCATTTTGGAGTCTGAATCAAT 3265  
 QY 3241 AGAGAAATTTAGACAACTATATCTGAGCTTATATATGATGATGATGATGATGATGATG 3300  
 DB 3266 AGAGAAATTTAGACAACTATATCTGAGCTTATATGATGATGATGATGATGATGATGATG 3325  
 QY 3301 AGTCTTAAGCGAAGAGACCTTTGAATATGCTGACTTGGAGACAGAACAGAGGTGCG 3360  
 DB 3326 AGTCTTAAGCGAAGAGACCTTTGAATATGCTGACTTGGAGACAGAACAGAGGTGCG 3385  
 QY 3361 ATGCAATTTGCGGTGTCATCAAGTCCAAACCTTCTGATCTCTGGAGCAGAGCTCCCTG 3420  
 DB 3386 ATGCAATTTGCGGTGTCATCAAGTCCAAACCTTCTGATCTCTGGAGCAGAGCTCCCTG 3445  
 QY 3421 GATATTAAGTTCAGAGCTATGACATGCTCTCACACTTCTTCAATTTGAGACTTCTGAG 3480  
 DB 3446 GATATTAAGTTCAGAGCTATGACATGCTCTCACACTTCTTCAATTTGAGACTTCTGAG 3505  
 QY 3481 GGAATCCCAATTAAGGTGCTAAGCAGGCAAGAAATAGCTTGGGTGTTTGCATCT 3540  
 DB 3506 GGAATCCCAATTAAGGTGCTAAGCAGGCAAGAAATAGCTTGGGTGTTTGCATCT 3565  
 QY 3541 ACTCAGGATACACATGCTGCTTTAAAGGCTCTGTCTGAAATTTGACAGCCCTTAATGAATACA 3600  
 DB 3566 ACTCAGGATACACATGCTGCTTTAAAGGCTCTGTCTGAAATTTGACAGCCCTTAATGAATACA 3625  
 QY 3601 GAAAGGCAAAATATCAAGTACGCTGACGAGGAGCTAGCTCACCAAGTCTCTTGGCTGTG 3660  
 DB 3626 GAAAGGCAAAATATCAAGTACGCTGACGAGGAGCTAGCTCACCAAGTCTCTTGGCTGTG 3685  
 QY 3661 GTACAGCCAAATGCAATTAATTTCCGAAATGCTTTGATTTGCTATTTGTGACCTC 3720  
 DB 3686 GTACAGCCAAATGCAATTAATTTCCGAAATGCTTTGATTTGCTATTTGTGACCTC 3745  
 QY 3721 AATGTTGATTAATTAAGTGAAGGCTCTGGGCTTCTTGAAGAGCAAGATCTTATCCAAAT 3780  
 DB 3746 AATGTTGATTAATTAAGTGAAGGCTCTGGGCTTCTTGAAGAGCAAGATCTTATCCAAAT 3805  
 QY 3781 CAAAGAGCTTTGATTTAGATGCTGTAAAGAAATTAAGATGATCTCAATCATGTG 3840  
 DB 3806 CAAAGAGCTTTGATTTAGATGCTGTAAAGAAATTAAGATGATCTCAATCATGTG 3865  
 QY 3841 GATTTGAATGTGTGACAAAGCTTTTGGGCGCGGTAGAGTGGGATGCTCTTATGGA 3900  
 DB 3866 GATTTGAATGTGTGACAAAGCTTTTGGGCGCGGTAGAGTGGGATGCTCTTATGGA 3925  
 QY 3901 GTTAACATTAATAAGGCTTTATGCTTACAGAAAGCAATTTCTGAGGAGACAGCTG 3960



**THIS PAGE BLANK (USPTO)**

```

Db 3926 GTTAACCTATTAGTGGCTTATGATGCTTGCCTTCAGAGCAATTTCTGTGAGGACAGCTG 3985
Qy 3961 AAGAAAGTGGAAATATGATCATGAGAACTCAACCTTATTAATTTCTGTAAATGAAGACC 4020
Db 3986 AAGAAAGTGGAAATATGATCATGAGAACTCAACCTTATTAATTTCTGTAAATGAAGACC 4045
Qy 4021 CAGTTTGTGTAAATATTTCTGTGTGAGAACTTAAAGTTTCAAAATACCAAGATGCT 4080
Db 4046 CAGTTTGTGTAAATATTTCTGTGTGAGAACTTAAAGTTTCAAAATACCAAGATGCT 4105
Qy 4081 TCAGTGTCCATAGAGATTAATTAATGAGCAAGAGACAGGCGGTGAGAAATTAACCTCT 4140
Db 4106 TCAGTGTCCATAGAGATTAATTAATGAGCAAGAGACAGGCGGTGAGAAATTAACCTCT 4165
Qy 4141 GAAGTGAAGCTGTCTCTGTGAGCTTGAAGATGATGAGAGAGAGAGAGAGAGAGAGAGAG 4200
Db 4166 GAAGTGAAGCTGTCTCTGTGAGAGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4225
Qy 4201 GATGAGAGCTTCAAGGCTCCATCATCATCTTCACTTCACTTCACTTCACTTCACTTCA 4260
Db 4226 GATGAGAGCTTCAAGGCTCCATCATCATCTTCACTTCACTTCACTTCACTTCACTTCA 4285
Qy 4261 CTGTACTTTATGAGAACTTGGCTGTGA 4287
Db 4286 CTGTACTTTATGAGAACTTGGCTGTGA 4312

```

## RESULT 2

AB079964  
ID AB079964 standard; cDNA; 4761 BP.

AC AB079964;

DT 23-DEC-2002 (first entry)

DE Human CD109 K1 protein encoding cDNA.

KM CD109; alpha2 macroglobulin; cholesteryl; cerebroprotective; cardiant;

KM immunosuppressive; haemostatic; anticoagulant; thrombolytic; human;

KM cardiovascular; vasotropic; gene therapy; CD109 K1; gene; 88.

OS Homo sapiens.

Key Location/Qualifiers

FT CDS 113..4450

FT /tag= a

FT /product= "CD109 K1"

PN MO200270696-A2.

PD 12-SEP-2002.

PF 07-MAR-2002; 2002W0-CA00292.

PR 07-MAR-2001; 2001US-273814P.

PA (SCHU/) SCHUH A.

PA (SUTR/) SUTHERLAND R D.

PI Schuh A, Sutherland RD;

DR MPI; 2002-713450/77.

XX P-PSDB; ABB82165.

XX New CD109 nucleic acids and polypeptides, useful in gene therapy,

XX particularly for treating strokes, myocardial infarctions, thrombosis,

XX chromocytopenia, autoimmune diseases, or organ or bone marrow

XX transplantation

XX Claim 1; Fig 1a; 156pp; English.

XX The invention relates to isolated nucleic acid molecules encoding CD109

CC polypeptides. These nucleic acid molecules include the human cDNA  
CC sequences comprising CD109 K1, CD109 K1-H, CD109 K1S or their variants.  
CC CD109 is a novel member of the alpha2 macroglobulin (alpha2M)/C3, C4, C5  
CC family of thioester-containing proteins. The CD109 polypeptides can be  
CC expressed by standard recombinant methodology. The CD109 nucleic acid,  
CC CD109 polypeptide or its mimetic, or the anti-CD109 antibody is useful  
CC for treating or detecting a disease or disorder, e.g. conditions  
CC associated with endothelial disease, platelet activation, activation  
CC of the coagulation or fibrinolytic systems, activation of T lymphocytes  
CC and of the complement system, quantitative or qualitative abnormalities  
CC of platelet function, increased or impaired platelet aggregation and  
CC activation, increased or impaired activation of the coagulation and/or  
CC fibrinolytic systems, or impaired or increased immune activation. These  
CC are also useful for treating cardiovascular disorders, stroke, myocardial  
CC infarction, thrombosis, embolism, peripheral vascular disease,  
CC thrombocytopenia, thrombocytopenia, autoimmune diseases, organ  
CC transplantation, or bone marrow transplantation. The CD109 nucleic acids  
CC or their antisense nucleotide sequence are useful in gene therapy for  
CC treating these conditions. The present sequence represents the human  
CC CD109 K1 cDNA sequence.

Sequence 4761 BP; 1420 A; 912 C; 1016 G; 1413 T; 0 other;

Query Match 98.5%; Score 4224.4; DB 24; Length 4761;  
Best Local Similarity 98.8%; Pred. No. 0;  
Matches 4286; Conservative 0; Mismatches 1; Indels 51; Gaps 1;

```

Qy 1 ATGCAAGGCGCCACCGCTCTGACCGCCGCGCCACCTCTGTGAGAGAGAGAGAGAGAGAG 60
Db 113 ATGCAAGGCGCCACCGCTCTGACCGCCGCGCCACCTCTGTGAGAGAGAGAGAGAGAGAG 172
Qy 61 GCCGTGCTCCCGGCGCTGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
Db 173 GCCGTGCTCCCGGCGCTGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 232
Qy 121 AATGACATATTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
Db 233 AATGACATATTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 292
Qy 181 GAGCTGCTCAAGACAGATCAAACTCACTGTCTGTCTGTGAGAGAGAGAGAGAGAGAGAG 240
Db 293 GAGCTGCTCAAGACAGATCAAACTCACTGTCTGTCTGTGAGAGAGAGAGAGAGAGAGAG 352
Qy 241 GAAAGAGCTCTTTAAGACATTAATCTTCATCACTCTGAAAGAGAGAGAGAGAGAGAG 300
Db 353 GAAAGAGCTCTTTAAGACATTAATCTTCATCACTCTGAAAGAGAGAGAGAGAGAGAGAG 412
Qy 301 ATTATGAGTACAGGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
Db 413 ATTATGAGTACAGGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 472
Qy 361 CGCTTATATTTGAGACCAAGAGATATCTGTTCATTAACAGAGAGAGAGAGAGAGAGAG 420
Db 473 CGCTTATATTTGAGACCAAGAGATATCTGTTCATTAACAGAGAGAGAGAGAGAGAGAG 532
Qy 421 AAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
Db 533 AAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 592
Qy 481 AAAAGCTCTTTAAACATTTCAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Db 593 AAAAGCTCTTTAAACATTTCAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 652
Qy 541 TCACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Db 653 TCACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 712
Qy 601 CTGAGTACTGTCTATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Db 713 CTGAGTACTGTCTATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Qy 661 GTTTCAGATATGATTAACAAATTTGAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720

```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

nucleic search, using sw model

January 16, 2004, 12:38:20 ; Search time 14986 Seconds  
(without alignments)  
11702.894 Million cell updates/sec

Sequence: 1 atcgagggccacgcctcct.....ttatggaacttgcgtgta 4287  
Sequence score: 4287  
US-10-020-095-3  
4287

Scoring table:  
IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database:

```

1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_by:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_hcg_hum:*
31: em_hcg_inh:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pla:*
35: em_hcg_rod:*
36: em_hcg_man:*
37: em_hcg_vtc:*
38: em_ey:*
39: em_hcg_hum:*
40: em_hcg_mus:*
41: em_hcg_other:*

```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Query Length	DB ID	Description
1	4224.4	98.5	4761	AX534934	AX534934 Sequence
2	4224.4	98.5	5883	AP410459	AP410459 Homo sapi
3	4224.4	98.5	5895	AX534938	AX534938 Sequence
4	4222.8	98.5	4761	AX534936	AX534936 Sequence
5	4222.8	98.5	5895	AX534940	AX534940 Sequence
6	4221.4	98.5	4335	AX537610	AX537610 Sequence
7	4219.8	98.4	4335	AX537612	AX537612 Sequence
8	3187.4	74.3	3535	AX534942	AX534942 Sequence
9	3185.8	74.3	3535	AX534944	AX534944 Sequence
10	2709.2	63.2	4735	AY083458	AY083458 Mus muscu
11	2709.2	63.2	5644	BC052443	BC052443 Mus muscu
12	2448.2	57.1	2938	AK095888	AK095888 Homo sapi
13	1969.4	45.9	2273	AK095888	AK095888 Homo sapi
14	730.8	17.0	1300	AX083146	AX083146 Sequence
15	232	5.4	163577	AL590428	AL590428 Homo sapi
16	230	5.1	176144	AC012408	AC012408 Homo sapi
17	219.6	5.1	5603	CIN43168	CIN43168 Homo sapi
18	203.8	4.8	4678	D83196	D83196 Homo sapi
19	182	4.2	251250	AC097023	AC097023 Rattus no
20	177.4	4.1	91419	AL591480	AL591480 Human DNA
21	175	4.1	149597	AC034271	AC034271 Homo sapi
22	174	4.1	4527	AX054954	AX054954 Sequence
23	174	4.1	4560	AX054954	AX054954 Sequence
24	171.6	4.0	2629	AY118302	AY118302 Drosophila
25	167.6	3.9	4715	GGGVOVM	GGGVOVM
26	167	3.9	4578	AY122084	AY122084 Drosophila
27	167	3.9	4578	DME269539	DME269539 Drosophila
28	160	3.7	144589	AC112668	AC112668 Mus muscu
29	160	3.7	191258	AC118254	AC118254 Mus muscu
30	155.4	3.6	4620	RNA113	RNA113
31	154.8	3.6	176144	AC012408	AC012408 Homo sapi
32	154.2	3.6	4636	RAT113A	RAT113A
33	151.2	3.5	4220	DME269538	DME269538
34	151.2	3.5	4575	GPMSPB	GPMSPB
35	149.8	3.5	2501	RAT113	RAT113
36	146.4	3.4	4950	XELENDO	XELENDO
37	143.8	3.4	122013	AC092399	AC092399 Drosophila
38	143.8	3.4	168414	AC092399	AC092399 Drosophila
39	143.8	3.4	168414	AC092399	AC092399 Drosophila
40	143.8	3.4	275699	AC115484	AC115484 Drosophila
41	143.8	3.4	337178	AE003649	AE003649 Drosophila
42	140.8	3.3	191258	AC118254	AC118254 Mus muscu
43	136	3.2	4623	BC036299	BC036299 Mus muscu
44	136	3.2	4672	BC051037	BC051037 Mus muscu
45	134.8	3.1	2689	AB026130	AB026130 Cypripus

## ALIGNMENTS

RESULT 1  
AX534934  
LOCUS AX534934 4761 bp DNA linear PAT 22-NOV-2002  
DEFINITION Sequence 1 from Patent WO02070696.  
ACCESSION AX534934  
VERSION AX534934.1 GI:25261474  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
REFERENCE  
1 Schuh, A. and Sutherland, R.D.  
TITLE Cdi09 nucleic acid molecules polypeptides and methods of use  
JOURNAL Patent: WO 02070696-A 1 12-SEP-2002;

**THIS PAGE BLANK (USPTO)**

Y	3670	ATGGCAGTAAATATTTGGCAAAATGTTGGATTGCTATTGGTCAATGTTGTA	3729
b	3833	ATGGCAGTAAATATTTGGCAAAATGTTGGATTGCTATTGGTCAATGTTGTA	3892
Y	3730	TATATATGAAAGGCTTGGGGCTTTCTAGAAAGCAAGATCTATCCAAATCAAGAACCC	3789
b	3893	TATATATGAAAGGCTTGGGGCTTTCTAGAAAGCAAGATCTATCCAAATCAAGAACCC	3952
Y	3790	TTTGATTTAGATGTTGCTGTGTAAGAAATTAAGATGATCTCAATCATGTGATTTGAT	3849
b	3953	TTTGATTTAGATGTTGCTGTGTAAGAAATTAAGATGATCTCAATCATGTGATTTGAT	4012
Y	3850	GTGTGTACAGCTTTTGGGGCCGGGTAGAGTGGCATGCTCTTATGGAAGTTACCTA	3909
b	4013	GTGTGTACAGCTTTTGGGGCCGGGTAGAGTGGCATGCTCTTATGGAAGTTACCTA	4072
Y	3910	TTAAGTGGCTTTATGAGTGGCTTCAAGAGCAATTTCTGTAGCGAGACAGTGAAGAAAGTG	3969
b	4073	TTAAGTGGCTTTATGAGTGGCTTCAAGAGCAATTTCTGTAGCGAGACAGTGAAGAAAGTG	4132
Y	3970	GAATATGATCATGAAAGTCAACCTCTATAGATCTGTAAATGAAGCAAGTTTGT	4029
b	4133	GAATATGATCATGAAAGTCAACCTCTATAGATCTGTAAATGAAGCAAGTTTGT	4192
Y	4030	GTTATATATCTGCTGTGAGAAATCTTAAAGTTTCAATATACCAAGATGCTTCAAGTCTC	4089
b	4193	GTTATATATCTGCTGTGAGAAATCTTAAAGTTTCAATATACCAAGATGCTTCAAGTCTC	4252
Y	4090	ATAGTGAATTAATCTATAGAGCCAGAGAGAGAGCGCGGTGAGAAAGTTCAACTGTGAAGTGAAG	4149
b	4253	ATAGTGAATTAATCTATAGAGCCAGAGAGAGAGCGCGGTGAGAAAGTTCAACTGTGAAGTGAAG	4312
Y	4150	CTGTCTCTCTGTGACCTTTGCAATGATGTCAGAGGCTGCGCTCTGTAGAGATGAGCT	4209
b	4313	CTGTCTCTCTGTGACCTTTGCAATGATGTCAGAGGCTGCGCTCTGTAGAGATGAGCT	4372
Y	4210	TCAGGCTCCCATCATGATCTTCACTCATTTTATTTTCTGTTCAAGCTTCTGTACTTT	4269
b	4373	TCAGGCTCCCATCATGATCTTCACTCATTTTATTTTCTGTTCAAGCTTCTGTACTTT	4432
Y	4270	ATGGAACCTTTGGCTGTGA	4287
b	4433	ATGGAACCTTTGGCTGTGA	4450

RESULT 2  
AF410459 5883 bp mRNA linear PRI 02-MAR-2002  
LOCUS Homo sapiens CD109 (CD109) mRNA, complete cds.  
DEFINITION AF410459  
ACCESSION AF410459  
VERSION AF410459.1 GI:19071288

KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
1 (bases 1 to 5883)  
Lin, M., Sutherland, D.R., Horefall, W., Toty, N., Yeo, E., Nayar, R., Wu, X.-F. and Schuch, A.C.  
Cell surface antigen CD109 is a novel member of the alpha(2) macroglobulin/C3, C4, C5 family of thioester-containing proteins  
Blood 99 (5), 1683-1691 (2002)

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
PAGES  
LOCATION  
1..5883  
/organism="Homo sapiens"

Submitted (14-AUG-2001) Medicine, University of Toronto, 1 King's College Circle, Room 7366, Toronto, Ontario M5S 1A8, Canada  
Location/Qualifiers  
1..5883  
/organism="Homo sapiens"

Y	1	ATGCAAGGCGCAACCGCTCTGACCGCGCGCAACCTCTCTGCTGTGACCGCGCGCTG	60
b	113	ATGCAAGGCGCAACCGCTCTGACCGCGCGCAACCTCTCTGCTGTGACCGCGCGCTG	172
Y	61	GCCGTGCTCCCGCGGCTCGGTTTCTGTGACAGCGCCAGGATCATCAGCGCCGAGGA	120
b	173	GCCGTGCTCCCGCGGCTCGGTTTCTGTGACAGCGCCAGGATCATCAGCGCCGAGGA	232
Y	121	AATGTGACTATTTGGGTGAGCTTCTGGAACCTGCGCTTCAAGAGTGTGAAGGCG	180
b	233	AATGTGACTATTTGGGTGAGCTTCTGGAACCTGCGCTTCAAGAGTGTGAAGGCG	292
Y	181	GAGTGTCAAGACACATCAAACTCACTGCTCTGTCTGTGAGACAGAGAGGCTTTT	240
b	293	GAGTGTCAAGACACATCAAACTCACTGCTCTGTCTGTGAGACAGAGAGGCTTTT	352
Y	241	GAAGAAAGGCTCTTTTAAGACACTTACTCTTCACTCACTCACTTGAAGAGTGAAG	300
b	353	GAAGAAAGGCTCTTTTAAGACACTTACTCTTCACTCACTCACTTGAAGAGTGAAG	412
Y	301	ATTATGAGCTAGCTGAACCGGACGACAGATGAGATTTATTTCTTAATGATACC	360

BASE COUNT 1697 a 1173 c 1241 g 1772 t  
ORIGIN  
3'UTR  
misc\_feature  
/gene="CD109"  
/note="alternative polyA site found on clone K1"  
Query Match 98.8%; Score 4224.4; DB 9; Length 5883;  
Best Local Similarity 98.8%; Pred. No. 0;  
Matches 4286; Conservative 0; Mismatches 1; Indels 51; Gaps 1;  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="K1"  
1..5883  
/gene="CD109"  
1..112  
/gene="CD109"  
113..4450  
/gene="CD109"  
/note="associated with the Gov alloantigen system"  
/product="CD109"  
/codon\_start=1  
/protein\_id="AA184159.1"  
/db\_xref="GI:19071209"  
/translation="MOGPPLLTAAHLCTAALAAVAPGRFLVTAPGIIIRPGNVIT  
GVELLEHCPQVYKAEILLKLTASNLTATSVLAEGVFEKGSFKTLTSLPLNSDELY  
ELRVYTGTOEILFNSSTRISFEKRSVLIQDPKALYKROKVPRIIVLFSFKPY  
KTSILNLIKDKPSNLIDQMLSDSDIGVISIKTROLSHPILGKSDIVQVNDOTYXOS  
FOVEEYVLPKEVYLOPLVCSMNSKHLNGLTAKYIYGRKVDIVLPLSPFGK  
KNITKTFKINGSANFSFNDKEMKNSNGSEYLDLSPGVEILITVTSVGT  
SRNVSTVFKROHDIIEFDYTVTLKPSLNFATVTRADGNQLTEERNNVIT  
VTORNTVEFGNSGNSGNOKMEAVOKINYTPQSGTFKEPILDESSSELQKAFILGS  
KSMANVSLKSPSKTYIQKTRDENIKGSPFLVVGSKRLKELSMYVRSQOLVA  
VGKONSTMSLTLPENSNTPKACVIVYIIEDEGELISVYLKIPQULVFNKIKIYMSVY  
KAEPSEKVSRLISVLOPDSIVGIVADKSVNLMAASVDITMENVHELELNTGYLIG  
MFNNSFAVFOECGLMLTDANLTQDYIDGVYDPSDITSWVAFVISEDGLTTPY  
SPHYRKPEETWMLDTNMGRIYOEFEVYDPSDITSWVAFVISEDGLTTPY  
ELQAPQPEFTFLNPSYVIGEEFALETITFNLYKATDEVVILIEKSDKFDILMTSNE  
INATGHOOTLIVPSEDATULFPIRPHILEIPTVYALASPTASDADATOMILVVAEG  
EKSTYSQSLDLDTNRLQSTLTKTSFSPFPRTVTSERVOITADIVLGPSINLASL  
IRMPYGGCEOMINPAPRIYILDYTLKKQLTDNLKEAUSFMRQYORELLYRDEG  
GSFAFYNDVPSGSTMSAFVLCFLPADPYIIDQNVHLRTYTLKQHSKNGEFMDP  
GVAHSELQGNKSPVTLIAYITSLGYRYQPNIDVOESIHFLSEFSGISDNYT  
LALITVYLSVSGSKAKALMLTMRABOEGMQFVWSBSKLSDSQPRSLDIEVAA  
YALISHPLOPOTSGIPIMRWLSRORNSLIGSFASDPTVTLKALSERPALMTERTN  
IOVTTTGSSSPYKRFILDTNRLILQTAELVAVOPAMVNSAAGFGAICOLNVVN  
VRAGSSRRRSRSTIONOEAFFDLVAVKEXKDLNHDVNCVCSFGSPSGMALMEVNL  
LSGFNVPEEALISLETYKVEYDHGKLNLYDSVNETOFCNITPAVRNFKXTNODAS  
VSIIDVYEPFRROAVSYNSVEVLSKSDLCSDVQCRPCDEGASGHHHSVIFLCFK  
LLYFELMLW"

**THIS PAGE BLANK (USPTO)**

Db 413 ATTATATAGCTAGTGTATACGAGAGTACCGAGATGAGATTTATTTCTTAATAGTACC 472  
Qy 361 CGCTTATCATTTTGTAGACCAAGAGATATCTGTCTTCAATGCAACAGAGCCCTTATAC 420  
Db 473 CGCTTATCATTTTGTAGACCAAGAGATATCTGTCTTCAATGCAACAGAGCCCTTATAC 532  
Qy 421 AAGCCAAAGCAAGAGTGAAGTTTGGCATTTGTAACACTCTTCTCAAGATTTTAAAGCTTAC 480  
Db 533 AAGCCAAAGCAAGAGTGAAGTTTGGCATTTGTAACACTCTTCTCAAGATTTTAAAGCTTAC 592  
Qy 481 AAAACCTCTTAAACATCTCTTAAGAGACCCCAATCAATTTGATCCACAGGTG 540  
Db 593 AAAACCTCTTAAACATCTCTTAAGAGACCCCAATCAATTTGATCCACAGGTG 652  
Qy 541 TCACAAAGAGTATCTTGAGTCAATTTCCAAAATTTTCAAGCATCTTCCCATCAATA 600  
Db 653 TCACAAAGAGTATCTTGAGTCAATTTCCAAAATTTTCAAGCATCTTCCCATCAATA 712  
Qy 601 CTGTGAGTCTGTCTATTCAGATTCAAGTGAATGACACACATATTAATCATATTTCAG 660  
Db 713 CTGTGAGTCTGTCTATTCAGATTCAAGTGAATGACACACATATTAATCATATTTCAG 772  
Qy 661 GTTTCAGAAATGATATTAACAAAATTTGAAAGTACATTCAGACACCATTAATTTGTTCT 720  
Db 773 GTTTCAGAAATGATATTAACAAAATTTGAAAGTACATTCAGACACCATTAATTTGTTCT 832  
Qy 721 ATGAATTTCTAAGATTTAAATGTGTACATCAAGCAAGATTAATGAGAAAGCAAGT 780  
Db 833 ATGAATTTCTAAGATTTAAATGTGTACATCAAGCAAGATTAATGAGAAAGCAAGT 892  
Qy 781 AAAGAGAGTAAAGCTTACATTTTACCTTTATCTTTTGGGAAAGAAATATTT 840  
Db 893 AAAGAGAGTAAAGCTTACATTTTACCTTTATCTTTTGGGAAAGAAATATTT 952  
Qy 841 ACAAACATTTAAGTAAATGATGTGCAAACTTCTTTTAAATGATGAAGATGAAA 900  
Db 953 ACAAACATTTAAGTAAATGATGTGCAAACTTCTTTTAAATGATGAAGATGAAA 1012  
Qy 901 AATGTATGATTTCTTCAAAATGAGATCTTGTGAATCTGGATCTATCTTCCCTGAGCA 960  
Db 1013 AATGTATGATTTCTTCAAAATGAGATCTTGTGAATCTGGATCTATCTTCCCTGAGCA 1072  
Qy 961 GTAGAAATTTTAAACAGTGAAGATCAAGTATTAAGAAATGTAAGCACT 1020  
Db 1073 GTAGAAATTTTAAACAGTGAAGATCAAGTATTAAGAAATGTAAGCACT 1132  
Qy 1021 AATGTGTTCTTCAACCAATGATTAATCATGATTTGATTTGATTAATTAATCTGCTTG 1080  
Db 1133 AATGTGTTCTTCAACCAATGATTAATCATGATTTGATTTGATTAATTAATCTGCTTG 1192  
Qy 1081 AAGCATCTCTCAACCTTCAACAGCTGTAAGATTAATCTGCTGATGAGCAACACTG 1140  
Db 1193 AAGCATCTCTCAACCTTCAACAGCTGTAAGATTAATCTGCTGATGAGCAACACTG 1252  
Qy 1141 ACTCTTGAAGAAAGAAATTAATGATGATTAATCAAGTGAACAGAGAACTTAATGAG 1200  
Db 1253 ACTCTTGAAGAAAGAAATTAATGATGATTAATCAAGTGAACAGAGAACTTAATGAG 1312  
Qy 1201 TACTGAGAGCGATCTAACAGTGAAGTAAATGAAAGTCTTCAAGAAATTAATTAAT 1260  
Db 1313 TACTGAGAGCGATCTAACAGTGAAGTAAATGAAAGTCTTCAAGAAATTAATTAAT 1372  
Qy 1261 ACTGTCCCCCAAGATGAACTTTTAAAGTGAATCCCAATCTGAGAGATTTCCAGTAG 1320  
Db 1373 ACTGTCCCCCAAGATGAACTTTTAAAGTGAATCCCAATCTGAGAGATTTCCAGTAG 1432  
Qy 1321 CTACAGTTGAAGGCTTATTTCTTGTAGTAAATGATGATGAGAGTCTATAGTCTGTTT 1380  
Db 1433 CTACAGTTGAAGGCTTATTTCTTGTAGTAAATGATGATGAGAGTCTATAGTCTGTTT 1492  
Qy 1381 AAGTCTCTGATGAAGTATGATCAATCAATTAATAACAGAGATGAATATTAAGGTGGA 1440

Db 1493 AAGTCTCTGATGAAGTATGATCAATCAATTAATAACAGAGATGAATATTAAGGTGGA 1552  
Qy 1441 TCGCCTTTTGAAGTGTGTTAGTGGCAACAAAGATTAAGAGATTAAGTGA 1500  
Db 1553 TCGCCTTTTGAAGTGTGTTAGTGGCAACAAAGATTAAGAGATTAAGTGA 1612  
Qy 1501 GTATCCAGGAGACAGTGTGCTGTAGAAAACAAATTCACAAATGTTCTTTAACTA 1560  
Db 1613 GTATCCAGGAGACAGTGTGCTGTAGAAAACAAATTCACAAATGTTCTTTAACTA 1672  
Qy 1561 CCAGAAATTTCTTGAAGTCCAAAAGCTGTGATATTTGTATTAATTAAGATGAG 1620  
Db 1673 CCAGAAATTTCTTGAAGTCCAAAAGCTGTGATATTTGTATTAATTAAGATGAG 1732  
Qy 1621 GAAATTAATGAAGTGTGTTAAATAATCTGTGACCTGTTTAAATTAAGATGAAG 1680  
Db 1733 GAAATTAATGAAGTGTGTTAAATAATCTGTGACCTGTTTAAATTAAGATGAAG 1792  
Qy 1681 CTATATTTGAAGTAAAGTGAAGCTGAACATCTGAAGAAAGTCTCTTGAATCTGTG 1740  
Db 1793 CTATATTTGAAGTAAAGTGAAGCTGAACATCTGAAGAAAGTCTCTTGAATCTGTG 1852  
Qy 1741 ACACAGCCTGACTCCATGTTGGGATTTGAGTGTGACAAAGTGTGATCTGATGAAT 1800  
Db 1853 ACACAGCCTGACTCCATGTTGGGATTTGAGTGTGACAAAGTGTGATCTGATGAAT 1912  
Qy 1801 GCCTTAATGATATTAACATGAGAAATGTGTCATGATGGAATTTAATAACAGGA 1860  
Db 1913 GCCTTAATGATATTAACATGAGAAATGTGTCATGATGGAATTTAATAACAGGA 1972  
Qy 1861 TATTAATTTAGGATTTCAATGATATCTTTGCAAGTCTTTAGAAATGTGACTGGGTA 1920  
Db 1973 TATTAATTTAGGATTTCAATGATATCTTTGCAAGTCTTTAGAAATGTGACTGGGTA 2032  
Qy 1921 TTGACAGATGCAAACTCCACAGAGATTAATTAATGATGTTTATGACATGACAGAT 1980  
Db 2033 TTGACAGATGCAAACTCCACAGAGATTAATTAATGATGTTTATGACATGACAGAT 2092  
Qy 1981 GCTGAGAGTTTATGAGAGAAATGAGACATATTTGAGATTAATCATGACTTTCTTGG 2040  
Db 2093 GCTGAGAGTTTATGAGAGAAATGAGACATATTTGAGATTAATCATGACTTTCTTGG 2152  
Qy 2041 GTTGAAGTCCACATGTCCGAAGATTTTCCAGAGCTTGAATTTGCTAGACCAAC 2100  
Db 2153 GTTGAAGTCCACATGTCCGAAGATTTTCCAGAGCTTGAATTTGCTAGACCAAC 2212  
Qy 2101 ATGGGTTACAGATTTTACCAAGATTTGAAGTAACTGAACCTGATTTCACTACTTCTGG 2160  
Db 2213 ATGGGTTACAGATTTTACCAAGATTTGAAGTAACTGAACCTGATTTCACTACTTCTGG 2272  
Qy 2161 GTGCTACTGTGTTTGTGATCTCTGAGAGCTGGGTCTTGAACCTAATCTACTCCAGTG 2220  
Db 2273 GTGCTACTGTGTTTGTGATCTCTGAGAGCTGGGTCTTGAACCTAATCTACTCCAGTG 2332  
Qy 2221 GAGTCCCAAGCTTCCACAGCTTTTCAATTTTGTGAATCTTCCCTACTCGTATACAGA 2280  
Db 2333 GAGTCCCAAGCTTCCACAGCTTTTCAATTTTGTGAATCTTCCCTACTCGTATACAGA 2392  
Qy 2281 GGTGAAGATTTGCTTGAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2340  
Db 2393 GGTGAAGATTTGCTTGAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2452  
Qy 2341 AAGGTATCAATGAAGAAAGTGAACAAATTTGATATTTCAATGACTTCAAGTGAATTAAT 2400  
Db 2453 AAGGTATCAATGAAGAAAGTGAACAAATTTGATATTTCAATGACTTCAAGTGAATTAAT 2512  
Qy 2401 GCCACAGGCCACAGAGAACCTTCTGTGTTCCAGTGAAGATGAGGCACTGTCTTTT 2460  
Db 2513 GCCACAGGCCACAGAGAACCTTCTGTGTTCCAGTGAAGATGAGGCACTGTCTTTT 2572  
Qy 2461 CCCATCAGGCCACATCTGGGAGAAATTTCTATCAAGTCAAGCTCTTTCAACCACT 2520  
Db 2573 CCCATCAGGCCACATCTGGGAGAAATTTCTATCAAGTCAAGCTCTTTCAACCACT 2632



**THIS PAGE BLANK (USPTO)**

2521 GCTTGTGATGCTGTCACCCAGATGATTTTAAAGCTGAGAGAAATGAAAAATCATAT 2580  
 2633 GCTTGTGATGCTGTCACCCAGATGATTTTAAAGCTGAGAGAAATGAAAAATCATAT 2692  
 2581 TCACAAATCATCTTATTAGACTGACGCAATAGGCTACAGAGTACCCGTAACCTTGG 2640  
 2693 TCACAAATCATCTTATTAGACTGACGCAATAGGCTACAGAGTACCCGTAACCTTGG 2752  
 2641 AGTTTCTCATTTCTCTCTAATACAGTGAATGCAAGTGAAGAGTTCAATCATCTGCAAT 2700  
 2753 AGTTTCTCATTTCTCTCTAATACAGTGAATGCAAGTGAAGAGTTCAATCATCTGCAAT 2812  
 2701 GGAAGATGTTCTGCTGCTTCCATGATGCTTACCTGCTGATTTGATGCTTATGCTG 2760  
 2813 GGAAGATGTTCTGCTGCTTCCATGATGCTTACCTGCTGATTTGATGCTTATGCTG 2872  
 2761 TGTGTGAAACAGAAACATGATTAATTTTGTCTCAATATTATTAATTTGATTTGCTGCT 2820  
 2873 TGTGTGAAACAGAAACATGATTAATTTTGTCTCAATATTATTAATTTGATTTGCTGCT 2932  
 2821 AAAAAAACAACCTGACAGATTAATTTGAAAGAAAAAGCTTTTCAATTAATGAGCAAGGT 2880  
 2933 AAAAAAACAACCTGACAGATTAATTTGAAAGAAAAAGCTTTTCAATTAATGAGCAAGGT 2992  
 2881 TACAGAGAACTTCTCTATGAGAGAAAGTGTCTTGTGAGCTTTTGGAGATTTAT 2940  
 2993 TACAGAGAACTTCTCTATGAGAGAAAGTGTCTTGTGAGCTTTTGGAGATTTAT 3052  
 2941 GACCTTCTGAGAGCACTTGTGTGCTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 3000  
 3053 GACCTTCTGAGAGCACTTGTGTGCTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 3112  
 3001 CCTTACATGATTAATGATGAGAAATGTTTACACAGAAACATTAACCTTGTGCTTAAAGCAAT 3060  
 3113 CCTTACATGATTAATGATGAGAAATGTTTACACAGAAACATTAACCTTGTGCTTAAAGCAAT 3172  
 3061 CAGAAATCCAAAGCTGATTTTGGGATCCAGAGAAAGTGTATATGATGAGCTTCAAGGT 3120  
 3173 CAGAAATCCAAAGCTGATTTTGGGATCCAGAGAAAGTGTATATGATGAGCTTCAAGGT 3232  
 3121 GGCATTAATAAGTCAAGTAACTTACAGCTATATTTGTAATCTTCTCTGAGGATTAAGA 3180  
 3233 GGCATTAATAAGTCAAGTAACTTACAGCTATATTTGTAATCTTCTCTGAGGATTAAGA 3292  
 3181 AAGTATCAGCTTACATGATGAGTGAAGTCTATTCATTTTGTGAGTCTGAATTCAGT 3240  
 3293 AAGTATCAGCTTACATGATGAGTGAAGTCTATTCATTTTGTGAGTCTGAATTCAGT 3352  
 3241 AAGAGATTTTGAAGATTAATTAATCTAGCCCTTATTAATTAATGATGAGTGAAG 3300  
 3353 AAGAGATTTTGAAGATTAATTAATCTAGCCCTTATTAATTAATGATGAGTGAAG 3412  
 3301 AGTCTTAAAGCAGAGAACTTTGATTAATGCTGATTTGAGAGAGAGAAACAGAGAGTGGC 3360  
 3413 AGTCTTAAAGCAGAGAACTTTGATTAATGCTGATTTGAGAGAGAGAAACAGAGAGTGGC 3472  
 3361 ATGCAATTTGAGTGTCTATGAGTCCAACTTTTGTGACTCTGAGAGCCAGCTCTCTG 3420  
 3473 ATGCAATTTGAGTGTCTATGAGTCCAACTTTTGTGACTCTGAGAGCCAGCTCTCTG 3532  
 3421 GATATTAAGTGTGAGCTTATGAGTGTCTGATGCTTCACTTCTTAATTAATTAATGAGT 3480  
 3533 GATATTAAGTGTGAGCTTATGAGTGTCTGATGCTTCACTTCTTAATTAATTAATGAGT 3592

3601 GAAAGACAAATATTCAGAGTACCGGTGACCGGGCTTACCTGACCAAGTCTT----- 3651  
 3713 GAAAGACAAATATTCAGAGTACCGGTGACCGGGCTTACCTGACCAAGTCTTGTAAAGTTT 3772  
 3652 -----CTTGTGTGGTACAGCA 3669  
 3773 CTGATTTGACACACACACACCGCTTACTCTTACAGACAGAGCTTGTGTGACAGCA 3832  
 3670 ATGCAATTAATTAATTTCCGCAATGTTTGTGATTTGATTTGTCAAGTCAATGTTGA 3729  
 3833 ATGCAATTAATTAATTTCCGCAATGTTTGTGATTTGATTTGTCAAGTCAATGTTGA 3892  
 3730 TTTAATGTAAGGCTTCTGAGTCTTCTGAGAGAGAGATTTCTCAAAATCAAGAGCC 3789  
 3893 TTTAATGTAAGGCTTCTGAGTCTTCTGAGAGAGAGATTTCTCAAAATCAAGAGCC 3952  
 3790 TTTGATTTGATGTTGCTGTTAAAGAAATTAAGTGAATCTCAATCATGTGATTTGAAT 3849  
 3953 TTTGATTTGATGTTGCTGTTAAAGAAATTAAGTGAATCTCAATCATGTGATTTGAAT 4012  
 3850 GTGTGTACAAAGCTTTTGGGCCCCGGGTAGAGTGTGATGCTTTATGGAATTAACCTA 3909  
 4013 GTGTGTACAAAGCTTTTGGGCCCCGGGTAGAGTGTGATGCTTTATGGAATTAACCTA 4072  
 3910 TTAAGTGTGTTTATGAGTCTTCAAGAGCAATTTCTGAGAGAGACAGTGAAGAAAGT 3969  
 4073 TTAAGTGTGTTTATGAGTCTTCAAGAGCAATTTCTGAGAGAGACAGTGAAGAAAGT 4132  
 3970 GAATATGATCATGAGAAACTCAACTTATTAGATTTCTGTAATGAAGAACCAAGTTTGT 4029  
 4133 GAATATGATCATGAGAAACTCAACTTATTAGATTTCTGTAATGAAGAACCAAGTTTGT 4192  
 4030 GTTAATATTCTGCTGTGAGAAACTTTAAAGTTTCAATTAACCAAGATGCTTCAAGTTC 4089  
 4193 GTTAATATTCTGCTGTGAGAAACTTTAAAGTTTCAATTAACCAAGATGCTTCAAGTTC 4252  
 4090 ATAGTGTATTAATTAAGACCAAGAGAGACAGCGGTGAAGATTAACCAAGTGAAG 4149  
 4253 ATAGTGTATTAATTAAGACCAAGAGAGACAGCGGTGAAGATTAACCAAGTGAAG 4312  
 4150 CTGTCTCTGCTGTGAGCTTTGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTG 4209  
 4313 CTGTCTCTGCTGTGAGCTTTGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTG 4372  
 4210 TCAGGCTCCATCATCACTCTTCAAGTCAATTTTATTTTCTGTTCAAGCTTGTACTTT 4269  
 4373 TCAGGCTCCATCATCACTCTTCAAGTCAATTTTATTTTCTGTTCAAGCTTGTACTTT 4432  
 4270 ATGGAACCTTGGCTGGA 4287  
 4433 ATGGAACCTTGGCTGGA 4450

RESULT 3  
 AX534938  
 LOCUS AX534938 5895 bp DNA linear PAT 22-NOV-2002  
 DEFINITION Sequence 5 from Patent WO02070696.  
 ACCESSION AX534938  
 VERSION AX534938.1 GI:25261482  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Placentalia; Catarrhini; Hominoidea; Homo.

## REFERENCE

1. Schuch, A. and Sutherland, R.D.  
 Cd10 nucleic acid molecules and methods of use  
 Patent: WO 02070696-A 5 12-SEP-2002;  
 Schmidt, Andre (Ed) Sutherland, Robert D. (CA)  
 Location/Qualifiers  
 1..5895  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"

Same as 1

**THIS PAGE BLANK (USPTO)**

QY	4150	CTCTCTCCCTGACCTTTCAGATGTCACAGGCTGCGCTCTTGTGAGATGAGAGCT	4209
DB	4313	CTGTCTCCCTGACCTTTCAGATGTCACAGGCTGCGCTCTTGTGAGATGAGAGCT	4372
QY	4210	TCAGGCTCCCATCATCATCTTTCAGATGTCACAGGCTGCGCTCTTGTGAGATGAGAGCT	4269
DB	4373	TCAGGCTCCCATCATCATCTTTCAGATGTCACAGGCTGCGCTCTTGTGAGATGAGAGCT	4432
QY	4270	ATGGAATTTGGCTGCTGCA	4287
DB	4433	ATGGAATTTGGCTGCTGCA	4450
RESULT 6			
LOCUS	AK537610	4335 bp	DNA
DEFINITION	Sequence 1 from Patent WO02070738.		linear
ACCESSION	AK537610		
VERSION	AK537610.1		GI:25269531
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Schuh, A. and Ouwehand, W.		
AUTHORS	Diagnosis and treatment of blood disorders		
TITLE	Patent: WO 02070738-A 1 12-SEP-2002;		
JOURNAL	Schuh, A. and Ouwehand, W.		
FEATURES			
source	1. .4335		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
	1. .>4335		
	/note="unlabeled protein product"		
	/codon_start=1		
	/protein_id="CAD57496.1"		
	/db_xref="GI:25269532"		
	/translation="MGGPPLTAAHLTCVCTALAAVAPGRFLVTPAGIIRPGANTYI		
	GVLELRCPQVTVKALKTASNTLVSVIAGVREKSPKTLPLPLNSADRIY		
	ELAVRTGRTDDELIFNSSTRLSFTKRIISVITDQKLYPKQKQVRIYITLSDPFPY		
	KISNITLILKOPKSNLIQOOSDGLVSLKTFQSLSHPLGMSIQOVNDQTYVOS		
	FOSEVYLPEKFTLQTPLYCSMNSKSLNGITAKYTGVPVGDVTLFLPLSFWGK		
	KKAITTKPKINGSANPSFNDENQWDSNGSLSEYLDSPGVELLITVTSVGI		
	SRVSTNVFTRKHDIIEFPDVTYVKSINFTATVVTADGNQVLEERRNNVIT		
	VTRNATYEVWSGNSGNQKBAVOKINITYTPSGTFKIEPILDESSELQKAYPLGS		
	KSNMAYSLFKPSKSTYIQLKTNDENIKVSPPELVISGKRLKEISYMWVSGOLVA		
	VGKONSTMFSLTPENSWTPKACVIVYIEDGELISDLKIPQVLYPKIKLYMSKV		
	KABSEKVSRLISVTPDSIVGIVAVDQVNLNANASNDITMENVHELKNTGYIYG		
	MEPNSFAVPOECGLMVLIDANLTQIDYIDGVDAEVAERMEENEGHIVIHDSLS		
	SPHTRKIPPEWTWILDTNMGRIYVFEFVTPDSITSMVATGVISDLGLITTV		
	ELQAFQPPFLPLNIPYVIRGEBRALKITLITVVALSPTAGDAVOMILVAESL		
	INATGHOQTLVPSDEDAIVLPPIRPHLGEIPITVVALSPTAGDAVOMILVAESL		
	EKSYOSILDLTNRLOSTLTKLTSFSPPTVTSERVOITAGDVLGSPINLASL		
	IRMPYGCQKMINFAPNIYILDYLTQKQVLTDLKELKALSPFQOGVORBELVPRDG		
	SFSFAGNVDPGSGTSLSAFVRLCELEDPYIDIDONVLRITTYLHLESEFSKIDNT		
	GRVHSELQGNKSPVTLTAVITSLIGRYKQPNIDVDSIHLESEFSKIDNT		
	LALITVALSSVSPKAKALMLTMRABEGMGQFWSSEKLSDSMQPSRLIDEVAA		
	YALLSHLQFQTSBGPIMTRLSRORNSPASTODTVALKALSEFALCOLNVN		
	IOVTVTGPSSSPVFLIDTHRLQLQRLAIVOPMAVNIISANGFSAICOLNVN		
	VYASGSSRRRSIQHQBAPDLDAVAKENDLHVDLAVCTSPSGSGMALEVN		
	LSGFWPSEALISLSTVKKVEYDHQKMLYLDVSVETQCVNIIPVRPKVSNODAS		
	VSIDVYBPRQAVNSNSEVLTSSCDLSDVQGRPCEDGASGSHHSVILFCFK		
	LIIFHEMLT"		

1296 a 837 c 922 g 1280 c

98.5%; Score 4221.4; DB 6; Length 4335;

Identity 98.8%; Pred. No. 0;

Negative 0; Mismatches 1; Indels 51; Gaps 1;

QY	1	ATGAGAGGCCCAACGCTCTGACCGCCGCCACCTCTCTGCTGTGACCGCCGCTTG	60
DB	1	ATGAGAGGCCCAACGCTCTGACCGCCGCCACCTCTCTGCTGTGACCGCCGCTTG	60
QY	61	GCCTGTGCTCCGCGCTCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG	120
DB	61	GCCTGTGCTCCGCGCTCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG	120
QY	121	ATGTGACTATGAGGCTGTGAGCTTCTGAAACCTGCTCTTCAAGGTGACTGTGAAGCG	180
DB	121	ATGTGACTATGAGGCTGTGAGCTTCTGAAACCTGCTCTTCAAGGTGACTGTGAAGCG	180
QY	181	GAGCTGCTCAAGACAGATCAAACTGCTCTTCTGCTGTGCTGTGCTGTGCTGTGCTGTG	240
DB	181	GAGCTGCTCAAGACAGATCAAACTGCTCTTCTGCTGTGCTGTGCTGTGCTGTGCTGTG	240
QY	241	GAAGAAAGCTCTTTTAAAGACCTTCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG	300
DB	241	GAAGAAAGCTCTTTTAAAGACCTTCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG	300
QY	301	ATTATGAGCTACGCTGTAACCGGACCTACCCAGATGAGATTTATTTCTTAATGATACC	360
DB	301	ATTATGAGCTACGCTGTAACCGGACCTACCCAGATGAGATTTATTTCTTAATGATACC	360
QY	361	CGCTTATCATTTGAGACCAAGAGATATCTGTCTTCAATTCAGACAGAGCCTTATAC	420
DB	361	CGCTTATCATTTGAGACCAAGAGATATCTGTCTTCAATTCAGACAGAGCCTTATAC	420
QY	421	AAGCCAAAGCAAGAGAGGTTGCGATTGTACCTCTTCTGATTTTAAAGCTTAC	480
DB	421	AAGCCAAAGCAAGAGAGGTTGCGATTGTACCTCTTCTGATTTTAAAGCTTAC	480
QY	481	AAAACCTTTTAAACATTTCTTAAAGGACCCCAATCAATTTGATCCAGCTGTTG	540
DB	481	AAAACCTTTTAAACATTTCTTAAAGGACCCCAATCAATTTGATCCAGCTGTTG	540
QY	541	TCACAAACAAAGTATCTTGAGAGCTATTTCCAAATCTTTCAGTATCTTCCATCAATA	600
DB	541	TCACAAACAAAGTATCTTGAGAGCTATTTCCAAATCTTTCAGTATCTTCCATCAATA	600
QY	601	CTTGTGAGCTGCTATTTCAAGTTCAAGTATGATGATGATGATGATGATGATGATGATG	660
DB	601	CTTGTGAGCTGCTATTTCAAGTTCAAGTATGATGATGATGATGATGATGATGATGATG	660
QY	661	GTTTCAGATATGATTTTACCAAAATTTGAAGTACTTTCGACAGCCATATATTTGTTCT	720
DB	661	GTTTCAGATATGATTTTACCAAAATTTGAAGTACTTTCGACAGCCATATATTTGTTCT	720
QY	721	ATGAATTTCAAGATTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATG	780
DB	721	ATGAATTTCAAGATTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATG	780
QY	781	AAAGAGAGCTAACGCTTACATTTTACCTTTTACCTTTTGGGAAAGAAAGAAATATT	840
DB	781	AAAGAGAGCTAACGCTTACATTTTACCTTTTACCTTTTGGGAAAGAAAGAAATATT	840
QY	841	ACAAAACATTTAAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATG	900
DB	841	ACAAAACATTTAAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATG	900
QY	901	AAATTAATGATTTTCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATG	960
DB	901	AAATTAATGATTTTCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATG	960
QY	961	GTAGAAATTTTAAACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGT	1020
DB	961	GTAGAAATTTTAAACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGT	1020
QY	1021	AATGTGTTCTTCAAGCAATGATTTATCATCTTGTGATTTTGAATATCTGCTGTTG	1080
DB	1021	AATGTGTTCTTCAAGCAATGATTTATCATCTTGTGATTTTGAATATCTGCTGTTG	1080
QY	1081	AAGCATCTTCAACCTTCAAGCACTGTGAAAGTAACTGCTGCTGATGCAACCAACTG	1140

**THIS PAGE BLANK (USPTO)**

Db	2161	GTGGCTACTGTTTGTGTGATCTCTGAGAGACCTGGGCTTGGAGTAAACAATCTACTCGAGTG	2220
Qy	2221	GAGCTCCAGACGCTTCCACCACTTTTTCATTTTGTGAATCTTCCCTACTCTGTATACAGA	2280
Db	2221	GAGCTCCAGACGCTTCCACCACTTTTTCATTTTGTGAATCTTCCCTACTCTGTATACAGA	2280
Qy	2281	GGTGGAAGATTTGCTTTGGAAATTAATCATATATTTGAAAGATGCACTGAGGTT	2340
Db	2281	GGTGGAAGATTTGCTTTGGAAATTAATCATATATTTGAAAGATGCACTGAGGTT	2340
Qy	2341	AAGGTAAATCATTTGAGAAAAGTACAAATTTGATTTCTATGACTTCAAGTGAATTAAT	2400
Db	2341	AAGGTAAATCATTTGAGAAAAGTACAAATTTGATTTCTATGACTTCAATTAATTAAT	2400
Qy	2401	GCCACAGGCGACACACAGACCCCTTGTGTTCCAGTGAAGATGGGCGAACCTGTTCTTTT	2460
Db	2401	GCCACAGGCGACACACAGACCCCTTGTGTTCCAGTGAAGATGGGCGAACCTGTTCTTTT	2460
Qy	2461	CCCATCAGGCCAACATCTGGGAGAAATTCCTATCAAGTCACAGCTCTTTCACCCACT	2520
Db	2461	CCCATCAGGCCAACATCTGGGAGAAATTCCTATCAAGTCACAGCTCTTTCACCCACT	2520
Qy	2521	GCTTCTGATGCTGTACCCAGATGATTTAGTAAAGCTGAAGAAATGAATAATCAAT	2580
Db	2521	GCTTCTGATGCTGTACCCAGATGATTTAGTAAAGCTGAAGAAATGAATAATCAAT	2580
Qy	2581	TCACAAATCAATCTTATTAAGACTTGAACATATGGCTACAGAGTACCCGTGAAAACCTTGG	2640
Db	2581	TCACAAATCAATCTTATTAAGACTTGAACATATGGCTACAGAGTACCCGTGAAAACCTTGG	2640
Qy	2641	AGTTTTCATTTTCCCTCTAATPACAGTACCTGGAGTGAAGAGTTCAGATCAGTCAATT	2700
Db	2641	AGTTTTCATTTTCCCTCTAATPACAGTACCTGGAGTGAAGAGTTCAGATCAGTCAATT	2700
Qy	2701	GGAGATGTTCTTGTCCTTCATCAATATGCTGACATGAGTGAAGAGTTCAGATCAGTCAATT	2760
Db	2701	GGAGATGTTCTTGTCCTTCATCAATATGCTGACATGAGTGAAGAGTTCAGATCAGTCAATT	2760
Qy	2761	TGAGTGGAACAGAAACATGATTAATTTTGTCTCCAAATATTTTCAATTTTGGATATCTGACT	2820
Db	2761	TGAGTGGAACAGAAACATGATTAATTTTGTCTCCAAATATTTTCAATTTTGGATATCTGACT	2820
Qy	2821	AAAAAGAAACAACGTGACATTAATTTGAAAGAAAAGCCTCTTCAATTTATGAGGCAAGT	2880
Db	2821	AAAAAGAAACAACGTGACATTAATTTGAAAGAAAAGCCTCTTCAATTTATGAGGCAAGT	2880
Qy	2881	TACCAAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTCAAGTCTTTGGGAATAT	2940
Db	2881	TACCAAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTCAAGTCTTTGGGAATAT	2940
Qy	2941	GACCCCTTCTGGAGACACTTGGTGTCACTTTTGTATTAAGATGTTTCTTTAGGCCGAT	3000
Db	2941	GACCCCTTCTGGAGACACTTGGTGTCACTTTTGTATTAAGATGTTTCTTTAGGCCGAT	3000
Qy	3001	CCTTACATAGATATTTGATCAGAAATGTTTACACAGACATACCTTGGCTTAAAGACAT	3060
Db	3001	CCTTACATAGATATTTGATCAGAAATGTTTACACAGACATACCTTGGCTTAAAGACAT	3060
Qy	3061	CAGAAATCCAAAGGTGAATTTTGGGATCCAGAAAGAGATTCATATGATGACTTCAAGGT	3120
Db	3061	CAGAAATCCAAAGGTGAATTTTGGGATCCAGAAAGAGATTCATATGATGACTTCAAGGT	3120
Qy	3121	GGCAATTAAGTCCAGTACACTTACAGCCTATATTTGTAATCTTCTCTGGATATACGA	3180
Db	3121	GGCAATTAAGTCCAGTACACTTACAGCCTATATTTGTAATCTTCTCTGGATATACGA	3180
Qy	3181	AAGATACAGCCTAACATGATGTCAGAAAGATCTATCAATTTTGGAGTCTGAATTCAGT	3240
Db	3181	AAGATACAGCCTAACATGATGTCAGAAAGATCTATCAATTTTGGAGTCTGAATTCAGT	3240
Qy	3241	AGAGGAATTTACAGCAATTAATCTAGCCCTTATACATTAATGATGTCATACAGTGGG	3300
Db	3241	AGAGGAATTTACAGCAATTAATCTAGCCCTTATACATTAATGATGTCATACAGTGGG	3300

**THIS PAGE BLANK (USPTO)**

QY	3301	AGTCTTAAGCGAAGAAAGCTTTGAAATGACGATTCGAGAGCGAACAAGAAAGTGGC	3350
Db	3301	AGTCTTAAGCGAAGAAAGCTTTGAAATGACGATTCGAGAGCGAACAAGAAAGTGGC	3350
QY	3351	ATGCATTTCTGGGTCTCATCAGAGTCCAACTTTCTGACTCTCGGCAAGCCAGCTCCCTG	3420
Db	3351	ATGCATTTCTGGGTCTCATCAGAGTCCAACTTTCTGACTCTCGGCAAGCCAGCTCCCTG	3420
QY	3421	GATATTGAAGTTGACGCTATGCACTGCTCTCAACTTCTTAACAATTTGACATTTCTGAG	3480
Db	3421	GATATTGAAGTTGACGCTATGCACTGCTCTCAACTTCTTAACAATTTGACATTTCTGAG	3480
QY	3481	GGAATCCCAATTATGAGGTGGCTTAAGCAGGCAAGAAATAGCTGGGGGTGTTTGACATCT	3540
Db	3481	GGAATCCCAATTATGAGGTGGCTTAAGCAGGCAAGAAATAGCTGGGGGTGTTTGACATCT	3540
QY	3541	ACTCAGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGACAGCCCTATGAAATACA	3600
Db	3541	ACTCAGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGACAGCCCTATGAAATACA	3600
QY	3601	GAAAGCAAAATATCCAGATGACCGTGAAGGGGCTAGCTCAACCAAGTCTT-----	3651
Db	3601	GAAAGCAAAATATCCAGATGACCGTGAAGGGGCTAGCTCAACCAAGTCTT-----	3651
QY	3652	-----CTTGGCTGGGTACAGCA 3659	
Db	3652	-----CTTGGCTGGGTACAGCA 3659	
QY	3661	CTGATTTGACACACAACCGCTACTCTCTTCAACAGCAGAGCTTGTCTGTGTGTAACGCA	3720
Db	3661	CTGATTTGACACACAACCGCTACTCTCTTCAACAGCAGAGCTTGTCTGTGTGTAACGCA	3720
QY	3670	ATGCGCATTAATATTTCCGCAATGCTTTTGAATTTGCTATTGTCAGCTCAATGTTGTA	3729
Db	3721	ATGCGCATTAATATTTCCGCAATGCTTTTGAATTTGCTATTGTCAGCTCAATGTTGTA	3780
QY	3720	TATAATGAAAGGCTTCTGGGCTCTTGAAGAAGAAATCATTCACAAATCAAGAAAGCC	3789
Db	3781	TATAATGAAAGGCTTCTGGGCTCTTGAAGAAGAAATCATTCACAAATCAAGAAAGCC	3840
QY	3790	TTTGAATTTAGATGTTGCTGTGAAGAAATAAAGATGATCTCAATCATGTGGAATTTGAAT	3849
Db	3841	TTTGAATTTAGATGTTGCTGTGAAGAAATAAAGATGATCTCAATCATGTGGAATTTGAAT	3900
QY	3850	GTTGTCTACAGCTTTTCCGGCCCGGGGTAGAGTGCGCATGCTCTTATGGAATTAACCTA	3909
Db	3901	GTTGTCTACAGCTTTTCCGGCCCGGGGTAGAGTGCGCATGCTCTTATGGAATTAACCTA	3960
QY	3910	TTAATGGCTTTATGAGTCTTCAAGAACAAATTTCTCTGAGCGAGACGTGAAGAAAGTG	3969
Db	3961	TTAATGGCTTTATGAGTCTTCAAGAACAAATTTCTCTGAGCGAGACGTGAAGAAAGTG	4020
QY	3970	GAAATGATCATGGAATACTCAACTCTATTTAGATTTCTGTAATGAAACCAGTTTTGT	4029
Db	4021	GAAATGATCATGGAATACTCAACTCTATTTAGATTTCTGTAATGAAACCAGTTTTGT	4080
QY	4030	GTTAATATTTCTGCTGTGAGAAACTTTAAAGTTTCAATACCAGATGCTTCAAGTCC	4089
Db	4081	GTTAATATTTCTGCTGTGAGAAACTTTAAAGTTTCAATACCAGATGCTTCAAGTCC	4140
QY	4090	ATTAGGATTAATCTATGAGCAAGAGGAGGCGGTGAAGTTCAACTCTGGAAGTAAG	4149
Db	4141	ATTAGGATTAATCTATGAGCAAGAGGAGGCGGTGAAGTTCAACTCTGGAAGTAAG	4200
QY	4250	CTGTCTCTCTGTGACCTTTGCAAGTATGTCAGAGGCTCCGCTCTTGTGAAGATGAGCT	4259
Db	4251	CTGTCTCTCTGTGACCTTTGCAAGTATGTCAGAGGCTCCGCTCTTGTGAAGATGAGCT	4260
QY	4260	CTGAGCTCCCATCATCACTCTTCAAGTCAATTTATTTCTGTTCAGACTTCTGTAATTT	4320
Db	4320	CTGAGCTCCCATCATCACTCTTCAAGTCAATTTATTTCTGTTCAGACTTCTGTAATTT	4320

RESULT 7

LOCUS	AX537612	4335 bp	DNA	linear	PAT 23-NOV-2002
DEFINITION	Sequence 3 from Patent WO02070738.				
ACCESSION	AX537612				
VERSION	AX537612.1 GI:25269535				
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 Schuh, A. and Ouehand, W. Diagnosis and treatment of blood disorders Patent: WO 02070738-A 3 12-SEP-2002; Schuh, Andre (CA)				
AUTHORS					
TITLE					
JOURNAL					
FEATURES	Location/Qualifiers				
source	1..4335 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" 1..>4335 /note="unnamed protein product" /codon_start=1 /protein_id="CAD57497.1" /db_xref="GI:25269535" /translation="MOGPPLLTAAHLCTALANAPRPRLVTARGIIRPGNVYII GVELLHPQSQVTKABLEKTKASNLTVSVLEAGVEFKSLTLPSPLENSADELY ELVETGKTDEILFNSSTRISFEPTKRIISVFIQDMLYKRGQVETKRIVLFSDPKY KTSILNLIKDPKSNLQQMSSQSDSLGSKTLPSSHPIGDMSSIOVQNDQTYQSS FOYSEVYLPEKFEVLTQPLKSSKSNLHNTITAKTYTGKPVAGDVLTFLPLSPFGK KNKIKTFKINGKSSFNDEEKKNVSSNGSLSEYLDSSPPVELITVTSVIGI SRVSTVNFKQKHDIYIEPDYTVKPLSNLPTATVTKRADNOLITLERRNVVIT VTQRNTEYMGSSNGKQKMEAVOKINITYTPQSGTKIERPILDESSLOLKAFLGS KSSMAVSHLKSFSKITYLOLTKIDENIKDGSPELVSSGMRKLEISYAVVSGOLVA VGSQNTSMFSLTPENSMTTPKACVIVYIIEBDBEIIISDVLKIPVLFKNIKLYKMSKY KAPSEKSVSLRIIVTQPDLSIVGIVADKSNLNNANSENIDIMEVNHLELNYGVYLG MNPNSFAVFOECGLAVLTDLNLTFRDYIDGVYDAAEYAEERPMENEGHIVIHQFSLGS SLAPHRKPEPTLIMLDNMGSRITYOEPEVPSISWATGVISDGLGLTTPV ELAOFRPFIEMLPYSVIRGEFALEITTPVILKATETKATETPIEKSDKDIIMTSE VTQRNTEYMGSSNGKQKMEAVOKINITYTPQSGTKIERPILDESSLOLKAFLGS KSSMAVSHLKSFSKITYLOLTKIDENIKDGSPELVSSGMRKLEISYAVVSGOLVA VGSQNTSMFSLTPENSMTTPKACVIVYIIEBDBEIIISDVLKIPVLFKNIKLYKMSKY KAPSEKSVSLRIIVTQPDLSIVGIVADKSNLNNANSENIDIMEVNHLELNYGVYLG MNPNSFAVFOECGLAVLTDLNLTFRDYIDGVYDAAEYAEERPMENEGHIVIHQFSLGS SLAPHRKPEPTLIMLDNMGSRITYOEPEVPSISWATGVISDGLGLTTPV ELAOFRPFIEMLPYSVIRGEFALEITTPVILKATETKATETPIEKSDKDIIMTSE VTQRNTEYMGSSNGKQKMEAVOKINITYTPQSGTKIERPILDESSLOLKAFLGS KSSMAVSHLKSFSKITYLOLTKIDENIKDGSPELVSSGMRKLEISYAVVSGOLVA VGSQNTSMFSLTPENSMTTPKACVIVYIIEBDBEIIISDVLKIPVLFKNIKLYKMSKY KAPSEKSVSLRIIVTQPDLSIVGIVADKSNLNNANSENIDIMEVNHLELNYGVYLG MNPNSFAVFOECGLAVLTDLNLTFRDYIDGVYDAAEYAEERPMENEGHIVIHQFSLGS SLAPHRKPEPTLIMLDNMGSRITYOEPEVPSISWATGVISDGLGLTTPV ELAOFRPFIEMLPYSVIRGEFALEITTPVILKATETKATETPIEKSDKDIIMTSE VTQRNTEYMGSSNGKQKMEAVOKINITYTPQSGTKIERPILDESSLOLKAFLGS KSSMAVSHLKSFSKITYLOLTKIDENIKDGSPELVSSGMRKLEISYAVVSGOLVA VGSQNTSMFSLTPENSMTTPKACVIVYIIEBDBEIIISDVLKIPVLFKNIKLYKMSKY KAPSEKSVSLRIIVTQPDLSIVGIVADKSNLNNANSENIDIMEVNHLELNYGVYLG MNPNSFAVFOECGLAVLTDLNLTFRDYIDGVYDAAEYAEERPMENEGHIVIHQFSLGS SLAPHRKPEPTLIMLDNMGSRITYOEPEVPSISWATGVISDGLGLTTPV ELAOFRPFIEMLPYSVIRGEFALEITTPVILKATETKATETPIEKSDKDIIMTSE VTQRNTEYMGSSNGKQKMEAVOKINITYTPQSGTKIERPILDESSLOLKAFLGS KSSMAVSHLKSFSKITYLOLTKIDENIKDGSPELVSSGMRKLEISYAVVSGOLVA VGSQNTSMFSLTPENSMTTPKACVIVYIIEBDBEIIISDVLKIPVLFKNIKLYKMSKY KAPSEKSVSLRIIVTQPDLSIVGIVADKSNLNNANSENIDIMEVNHLELNYGVYLG MNPNSFAVFOECGLAVLTDLNLTFRDYIDGVYDAAEYAEERPMENEGHIVIHQFSLGS SLAPHRKPEPTLIMLDNMGSRITYOEPEVPSISWATGVISDGLGLTTPV ELAOFRPFIEMLPYSVIRGEFALEITTPVILKATETKATETPIEKSDKDIIMTSE VTQRNTEYMGSSNGKQKMEAVOKINITYTPQSGTKIERPILDESSLOLKAFLGS KSSMAVSHLKSFSKITYLOLTKIDENIKDGSPELVSSGMRKLEISYAVVSGOLVA VGSQNTSMFSLTPENSMTTPKACVIVYIIEBDBEIIISDVLKIPVLFKNIKLYKMSKY KAPSEKSVSLRIIVTQPDLSIVGIVADKSNLNNANSENIDIMEVNHLELNYGVYLG MNPNSFAVFOECGLAVLTDLNLTFRDYIDGVYDAAEYAEERPMENEGHIVIHQFSLGS SLAPHRKPEPTLIMLDNMGSRITYOEPEVPSISWATGVISDGLGLTTPV ELAOFRPFIEMLPYSVIRGEFALEITTPVILKATETKATETPIEKSDKDIIMTSE VTQRNTEYMGSSNGKQKMEAVOKINITYTPQSGTKIERPILDESSLOLKAFLGS KSSMAVSHLKSFSKITYLOLTKIDENIKDGSPELVSSGMRKLEISYAVVSGOLVA VGSQNTSMFSLTPENSMTTPKACVIVYIIEBDBEIIISDVLKIPVLFKNIKLYKMSKY KAPSEKSVSLRIIVTQPDLSIVGIVADKSNLNNANSENIDIMEVNHLELNYGVYLG MNPNSFAVFOECGLAVLTDLNLTFRDYIDGVYDAAEYAEERPMENEGHIVIHQFSLGS SLAPHRKPEPTLIMLDNMGSRITYOEPEVPSISWATGVISDGLGLTTPV ELAOFRPFIEMLPYSVIRGEFALEITTPVILKATETKATETPIEKSDKDIIMTSE VTQRNTEYMGSSNGKQKMEAVOKINITYTPQSGTKIERPILDESSLOLKAFLGS KSSMAVSHLKSFSKITYLOLTKIDENIKDGSPELVSSGMRKLEISYAVVSGOLVA VGSQNTSMFSLTPENSMTTPKACVIVYIIEBDBEIIISDVLKIPVLFKNIKLYKMSKY KAPSEKSVSLRIIVTQPDLSIVGIVADKSNLNNANSENIDIMEVNHLELNYGVYLG MNPNSFAVFOECGLAVLTDLNLTFRDYIDGVYDAAEYAEERPMENEGHIVIHQFSLGS SLAPHRKPEPTLIMLDNMGSRITYOEPEVPSISWATGVISDGLGLTTPV ELAOFRPFIEMLPYSVIRGEFALEITTPVILKATETKATETPIEKSDKDIIMTSE VTQRNTEYMGSSNGKQKMEAVOKINITYTPQSGTKIERPILDESSLOLKAFLGS KSSMAVSHLKSFSKITYLOLTKIDENIKDGSPELVSSGMRKLEISYAVVSGOLVA VGSQNTSMFSLTPENSMTTPKACVIVYIIEBDBEIIISDVLKIPVLFKNIKLYKMSKY KAPSEKSVSLRIIVTQPDLSIVGIVADKSNLNNANSENIDIMEVNHLELNYGVYLG MNPNSFAVFOECGLAVLTDLNLTFRDYIDGVYDAAEYAEERPMENEGHIVIHQFSLGS SLAPHRKPEPTLIMLDNMGSRITYOEPEVPSISWATGVISDGLGLTTPV ELAOFRPFIEMLPYSVIRGEFALEITTPVILKATETKATETPIEKSDKDIIMTSE VTQRNTEYMGSSNGKQKMEAVOKINITYTPQSGTKIERPILDESSLOLKAFLGS KSSMAVSHLKSFSKITYLOLTKIDENIKDGSPELVSSGMRKLEISYAVVSGOLVA VGSQNTSMFSLTPENSMTTPKACVIVYIIEBDBEIIISDVLKIPVLFKNIKLYKMSKY KAPSEKSVSLRIIVTQPDLSIVGIVADKSNLNNANSENIDIMEVNHLELNYGVYLG MNPNSFAVFOECGLAVLTDLNLTFRDYIDGVYDAAEYAEERPMENEGHIVIHQFSLGS SLAPHRKPEPTLIMLDNMGSRITYOEPEVPSISWATGVISDGLGLTTPV ELAOFRPFIEMLPYSVIRGEFALEITTPVILKATETKATETPIEKSDKDIIMTSE VTQRNTEYMGSSNGKQKMEAVOKINITYTPQSGTKIERPILDESSLOLKAFLGS KSSMAVSHLKSFSKITYLOLTKIDENIKDGSPELVSSGMRKLEISYAVVSGOLVA VGSQNTSMFSLTPENSMTTPKACVIVYIIEBDBEIIISDVLKIPVLFKNIKLYKMSKY KAPSEKSVSLRIIVTQPDLSIVGIVADKSNLNNANSENIDIMEVNHLELNYGVYLG MNPNSFAVFOECGLAVLTDLNLTFRDYIDGVYDAAEYAEERPMENEGHIVIHQFSLGS SLAPHRKPEPTLIMLDNMGSRITYOEPEVPSISWATGVISDGLGLTTPV ELAOFRPFIEMLPYSVIRGEFALEITTPVILKATETKATETPIEKSDKDIIMTSE VTQRNTEYMGSSNGKQKMEAVOKINITYTPQSGTKIERPILDESSLOLKAFLGS KSSMAVSHLKSFSKITYLOLTKIDENIKDGSPELVSSGMRKLEISYAVVSGOLVA VGSQNTSMFSLTPENSMTTPKACVIVYIIEBDBEIIISDVLKIPVLFKNIKLYKMSKY KAPSEKSVSLRIIVTQPDLSIVGIVADKSNLNNANSENIDIMEVNHLELNYGVYLG MNPNSFAVFOECGLAVLTDLNLTFRDYIDGVYDAAEYAEERPMENEGHIVIHQFSLGS SLAPHRKPEPTLIMLDNMGSRITYOEPEVPSISWATGVISDGLGLTTPV ELAOFRPFIEMLPYSVIRGEFALEITTPVILKATETKATETPIEKSDKDIIMTSE VTQRNTEYMGSSNGKQKMEAVOKINITYTPQSGTKIERPILDESSLOLKAFLGS KSSMAVSHLKSFSKITYLOLTKIDENIKDGSPELVSSGMRKLEISYAVVSGOLVA VGSQNTSMFSLTPENSMTTPKACVIVYIIEBDBEIIISDVLKIPVLFKNIKLYKMSKY KAPSEKSVSLRIIVTQPDLSIVGIVADKSNLNNANSENIDIMEVNHLELNYGVYLG MNPNSFAVFOECGLAVLTDLNLTFRDYIDGVYDAAEYAEERPMENEGHIVIHQFSLGS SLAPHRKPEPTLIMLDNMGSRITYOEPEVPSISWATGVISDGLGLTTPV ELAOFRPFIEMLPYSVIRGEFALEITTPVILKATETKATETPIEKSDKDIIMTSE VTQRNTEYMGSSNGKQKMEAVOKINITYTPQSGTKIERPILDESSLOLKAFLGS KSSMAVSHLKSFSKITYLOLTKIDENIKDGSPELVSSGMRKLEISYAVVSGOLVA VGSQNTSMFSLTPENSMTTPKACVIVYIIEBDBEIIISDVLKIPVLFKNIKLYKMSKY KAPSEKSVSLRIIVTQPDLSIVGIVADKSNLNNANSENIDIMEVNHLELNYGVYLG MNPNSFAVFOECGLAVLTDLNLTFRDYIDGVYDAAEYAEERPMENEGHIVIHQFSLGS SLAPHRKPEPTLIMLDNMGSRITYOEPEVPSISWATGVISDGLGLTTPV ELAOFRPFIEMLPYSVIRGEFALEITTPVILKATETKATETPIEKSDKDIIMTSE VTQR				



**THIS PAGE BLANK (USPTO)**



Creation date: 02-19-2004  
Indexing Officer: AGOMEZ - ALFREDO GOMEZ, JR.  
Team: OIPEBackFileIndexing  
Dossier: 09915515

Legal Date: 01-05-2004

No.	Doccode	Number of pages
1	SRNT	64

Total number of pages: 64

Remarks:

Order of re-scan issued on .....

